AD_____

Award Number: W81XWH-10-1-0463

TITLE: Origins of DNA Replication and Amplification in the Breast Cancer Genome

PRINCIPAL INVESTIGATOR: Susan A. Gerbi, Ph.D.
Alexander Brodsky, Ph.D.
Ben Raphael, Ph.D.

CONTRACTING ORGANIZATION: Brown University
Providence, RI 02912

REPORT DATE: September 2013

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2013	Final	1 September 2010 – 31 August 2013

4. TITLE AND SUBTITLE

Origins of DNA Replication and Amplification in the Breast Cancer Genome

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0463

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Susan A. Gerbi
Alexander Brodsky
Ben Raphael
E-Mail: Susan_Gerbi@Brown.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Brown University
Providence, RI 02912

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The major goal of this DOD grant was to test whether a correlation exists between sites of DNA amplification and estrogen receptor (ER) binding in the breast cancer genome. Correlations would support our hypothesis that ER adjacent to replication origins may interact with the replication machinery to drive DNA amplification, a hallmark of cancers. Our Specific Aims were:
(1) Map replication origins in the MCF-7 breast cancer genome by genomic sequencing
(2) Compare the replication origin maps between breast cancer and normal breast cells (MCF-10A)
(3) Correlate the origin map data with sites of DNA amplification and estrogen receptor binding If time remains, do pilot runs to map replication origins in ER+ human breast cancer tissue and sites of DNA amplification. For these aims, we have developed and refined the method of nascent strand sequencing (NS-Seq). This method uses lambda exonuclease, a method we developed to enrich newly synthesized DNA. There have been several reports of replication origin mapping in metazoan origins, but their results do not agree well. In our DOD-funded research, we discovered that lambda exonuclease has biases, resulting in contamination of the nascent strand DNA preparations. We developed bias corrections to apply NS-Seq to MCF-7 and MCF-10A cells.

15. SUBJECT TERMS
estrogen receptor, DNA amplification, replication origins

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	151	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

FINAL PROGRESS REPORT

INTRODUCTION

Genetic instability and rearrangements, including gene amplification, is a hallmark of many cancers (Santarius et al. 2010). Gene amplification is common in breast cancer genomes (McBride et al. 2012). Amplification of the HER2 (ErbB2/Neu) gene, which encodes human epidermal growth factor 2, occurs in ~25% of invasive breast cancers and in 50-60% of ductal carcinoma *in situ.* HER2 amplification and concomitant over-expression of this growth factor promotes cancer cell growth in a variety of tissue environments, acting as a metastasis-promoting factor. There are many examples of DNA amplification of other oncogenes as well. DNA amplification plays a role in establishing the malignant cell phenotype in cancer (Nikolsky et al., 2008). It would be desirable to prevent gene amplification, thereby moderating the aggressive growth of breast cancer cells. The problem is that no one knows what triggers gene amplification. Our current research in a model system suggests that the trigger may be a transcription factor such as the receptor for the steroid hormone estrogen. Cancer is believed to occur after a build-up of somatic mutations or other genomic changes. We wish to ask if a genomic change (genetic or epigenetic) might produce novel binding site(s) for the estrogen receptor (ER) near a replication origin to cause re-replication. The novel ER binding site would be absent in normal cells prior to amplification. Our hypothesis, based on our recent data from a model system, is that the ER interacts with the replication machinery to drive re-replication, resulting in amplification.

The basic mechanism underlying DNA amplification has not yet been elucidated but the current dogma is that it is a result of DNA double strand breaks (Debatisse and Malfoy, 2005). The earlier model of nested replication forks in an onionskin structure leading to DNA amplification (Stark and Wahl, 1984) has never been disproven, but fell into disfavor due to lack of supporting evidence. Recently, evidence has been mounting that DNA re-replication may indeed lead to DNA amplification. Several mechanisms have been proposed to explain how DNA replication may give rise to DNA amplification (Payen et al. 2008; Brewer et al., 2011; Finn and Li, 2013). Increases in the pre-replication complex proteins MCMs and Cdt1 have been shown to induce DNA amplification in yeast (Gopalakrishnan et al., 2001; Nguyen et al., 2001; Green et al., 2006 and 2010; Finn and Li, 2013) and increased Cdt1 results in re-replication in human cells (Dorn et al., 2000). The N-terminus of Cdt1 is important for re-replication, perhaps through interactions with PCNA and/or cyclin (Teer and Dutta, 2008). Cdt1 and its inhibitor geminin are deregulated in human tumors (Petropoulou et al., 2008) and depletion of geminin leads to re-replication (Zhu et al., 2004). In addition, over-expression of pre-replication complex proteins is associated with cancer progression (Ren et al., 2006; Liontis et al., 2007). Moreover, stalled replication forks (such as in the onionskin structure) and DNA re-replication lead to DNA breakage and rearrangements (Green and Li, 2005; Raveendranathan et al., 2006; Zhu and Dutta, 2006; Dutta, 2007; Hook et al. 2007) which is a hallmark of cancer. DNA amplification may be more common than originally thought and has been reported to occur even in normal cells (Gomez 2008; Gomez and Antequera 2008; Dorn et al., 2009).

The trigger for DNA amplification has not been identified, and we hypothesize that it may be transcription factors such as the estrogen receptor. There appears to be a link between the steroid hormone estrogen and many forms of breast cancer, but the detailed mechanism is unknown. Estrogen can turn on gene expression and thus activate the production of the proteins encoded by these genes. Our recent results in a model system indicated that a steroid hormone can induce gene amplification in which re-replication creates extra copies of the gene. This in turn will also increase production of the protein encoded by the amplified gene. Our data in a model system suggest that the hormone receptor binds to DNA that is adjacent to the origin replication complex (ORC); ORC serves as a landing pad for other components of the pre-replication complex. Hormonal induction of gene amplification is a new paradigm for how hormones work, and we wish to see if it applies to breast cancer. We wish to examine if a correlation exists between estrogen receptor (ER) binding at novel sites in the breast cancer genome and juxtaposition with replication origins that escape normal cellular controls and re-replicate, leading to DNA amplification. The observations that estrogen induces cell proliferation by retention of MCM proteins in the nucleus and by induction of the loading factor Cdt1 (Pan et al., 2006) support our hypothesis. Similarly, the finding that the transcription factor c-Myc interacts with the pre-replication complex to control DNA replication (Dominguez-Sola et al., 2007; Lebofsky and Walter, 2007) and that the androgen receptor interacts with MCM7 of the pre-replication complex (Shi, 2008) provides precedence for our hypothesis that the ligand-bound estrogen receptor may play a direct role in regulating replication origins beyond its traditional role as a transcription factor.

Our research may provide a new paradigm for hormonal induction of breast cancer via gene amplification, leading to new methods of diagnosis and treatment.

BODY

In the research supported by this grant, we proposed to map origins of DNA replication in the genome of breast cancer cells (MCF-7) and normal breast cells (MCF-10A). Next, we planned to identify the subset of replication origins that reside in areas of DNA amplification in the breast cancer genome and ask what sequences at the origins in regions of DNA amplification may distinguish them from origins in non-amplified regions of the breast cancer genome. Furthermore, the sequences from origins in regions of DNA amplification in breast cancer cells can be compared to the same origins in normal breast cells to explore what changes might have occurred in the DNA sequence of these origins. One such change could be the proximity of estrogen receptor binding sites near regions from areas of DNA amplification. We planned to test this possibility by comparing estrogen receptor ChIP-Seq data to our replication origin maps. All of these plans hinge upon identification of DNA replication origins genome-wide, and we have expended much effort on this.

Our **Specific Aims** and the study design listed for this grant were:

(1) Map replication origins in the human genome. Using MCF-7 breast cancer cells, we proposed to continue our development of methodology to map replication

origins in the human genome by sequencing short newly synthesized (nascent) strands that derive from sites encompassing replication origins. Our approach uses lambda-exonuclease digestion to enrich for nascent DNA, a method that we have pioneered.

(2) Comparison of replication origin maps between breast cancer and normal breast cells. Having developed the replication origin mapping technology in Specific Aim (1), we proposed to apply it to breast cancer cells (MCF-7) that are estrogen receptor positive (ER+) and compare these data to normal breast epithelial cells (MCF-10A). These results will indicate if replication origin usage changes between normal and breast cancer cells. Are new replication origins activated in breast cancer cells as compared to normal cells?

(3) Correlation of origin map data with sites of (a) DNA amplification and (b) estrogen receptor binding. Regions of DNA amplification in the cell lines to be used have already been determined by comparative genome hybridization (CGH)(Neve et al. 2006) and will be confirmed by our genome wide sequencing data. Use of these data will allow us to identify which replication origins are potentially origins of re-replication, resulting in DNA amplification. We want to know what triggered these origins to re-replicate. We will analyze whether a correlation exists between origins of re-replication and ER binding sites. These results will support or refute the hypothesis that ER may bind next to the replication machinery and induce DNA amplification.

The results from the proposed experiments will serve as the foundation for comparable experiments in surgically derived breast cancer tissue. These experiments are beyond the time frame of this grant, but we are already stockpiling tissue samples for these future experiments. Our proposed study could provide a new paradigm for hormonal induction of breast cancer via gene amplification, leading to new methods of diagnosis and treatment.

The text that follows is the final report for this grant, organized according the Statement of Work that was listed in the approved grant application which was organized according to three Specific Aims above. Figures and tables are located in the Supporting Data at the end of the document and in Appendix 1 and Appendix 2.

Description of Research Results

Task (1) Map replication origins in the human genome.
Subtask (1a) (months 1-6) – preparation of short nascent DNA from MCF-7 cells.

The Brodsky lab grew MCF-7 cells to mid-log stage and gave them to the Gerbi lab (postdoc Michael Foulk) for DNA isolation. The nascent strand sequencing work flow is as follows:

We developed methodology using MCF-7 breast cancer cells to derive a genomic map of replication origins by Illumina sequencing of short nascent strands. Postdoc Michael Foulk in the Gerbi lab prepared nascent strands from MCF-7 cells as follows:

Nascent-strand-Seq Work Flow (Foulk method)

- Prepare genomic DNA from asynchronous MCF-7 cells with DNAzol (~30% in S-phase)
- Purify Replicative Intermediate (RI) DNA on BND-cellulose (100 ug input)
- Phosphorylate ends with T4-polynucleotide kinase
- Enrich for Nascent Strands by digesting with lambda-exonuclease (λ-exonuclease = Lexo)
- Size select for 1 kb-2 kb nascent strands on low melting point agarose gel to eliminate Okazaki fragments
- Test enrichment of the MYC origin of replication by Real-Time PCR

This was the method we used for the MCF-7 nascent strands used for sequencing on the Illumina GAIIX platform. Subsequently, the same DNA preparation was used for the first run on the Illumina Hi-Seq platform. Real time PCR revealed that this nascent strand DNA preparation had 45-fold enrichment when tested for the Myc origin of DNA replication. Later nascent strand DNA preparations followed the method of Cadoret et al. (2008; personal communication) that is similar to the recent method of Cayrou et al. (2012b) and replaces the BND-cellulose column with a sucrose gradient; this method gave even higher levels of origin enrichment. The steps for the Cadoret method are summarized below:

Nascent-strand-Seq Work Flow (Cadoret method)

- Prepare genomic DNA from asynchronous MCF-7 cells with DNAzol (~30% in S-phase)
- Heat denature the DNA (95^0C for 5 min) and run on a 5-30% sucrose gradient; select the fractions with 300-2500 nt DNA.
- Phosphorylate ends with T4-polynucleotide kinase
- Enrich for Nascent Strands by digesting with lambda-exonuclease (λ-exonuclease = Lexo)
- Size select for 1 kb-2 kb nascent strands on low melting point agarose gel to eliminate Okazaki fragments
- Test enrichment of the Myc origin of replication by quantitative PCR

With either method, after enrichment of the MCF-7 nascent strand DNA, it was then subjected to DNA sequencing as follows:

- Subject nascent strands to fragmentation followed by making them double stranded with random primers and Klenow
- Standard library preparation for Illumina sequencing (200-500bp fragments)
- Sequence library on the Illumina GAIIx platform (pilot experiment) or High-Seq (subsequent experiments) using 42 bp single end reads
- Filter and align reads to the human genome (build hg18; MCF-7 break points: Hampton et al., 2008) and call peaks (using genomic input DNA as normalization control)

Several labs are now using the methodology we developed for nascent strand sequencing (NS-Seq) whose basis resides in the use of lambda exonuclease to enrich

nascent strands coupled with size selection. Our NS-Seq protocol is based on our earlier report (Gerbi and Bielinsky, 1997; Bielinsky and Gerbi, 1998) that nascent DNA is resistant to lambda exonuclease digestion because of the presence of a 5' RNA primer. This allows the parental DNA to be digested while the nascent DNA is untouched. The nascent strands were size selected on gels for 1-2 kb, which gave greater origin enrichment than a 0.5-1 kb fraction that may have Okazaki fragment contamination. Using the c-Myc origin to assess for enrichment, the average of the several preparations used for sequencing had 45-fold or greater enrichment of nascent strands. Interestingly, in assessing this enrichment at the c-Myc origin, we discovered that the preferred origin resided in the second exon of the gene while it was previously determined to reside in the promoter of the gene (in HeLa cells: Tao et al., 2000). This observation was confirmed in our NS-Seq data, suggesting plasticity of origin usage at the c-Myc gene in different cell types.

Sub-task (1b) (months 7-8) – sequencing and analysis of results to map replication origins in the MCF-7 human breast cancer cell genome.

As reported in the year one progress report, we switched from Helicos sequencing to Illumina sequencing, and we reported our first run of MCF-7 nascent strand DNA on the Illumina GAIIx machine. Since that time, we have used the more powerful Illumina Hi-Seq2000 machine as the platform for sequencing nascent strands. The new results confirm and extend our earlier data from the Illumina GAIIx machine. Of the three samples run on the Illumina Hi-Seq2000, the first sample was derived from the same material that had been used for the Illumina GAIIx machine. The other two samples were from different preparations of nascent strands, providing biological replicates.

	# reads mapped:	% of total reads from that run
GAIIx:	11,805,186 mapped	~44.4% of total reads
Lane1:	54,642,610 mapped	~42.61% of total reads
Lane3:	84,037,782 mapped	~60.32% of total reads
Lane5:	89,097,569 mapped	~65.4% of total reads
Total:	**239,583,147 NS reads mapped**	

Input: 179,965,523 input reads mapped ~92.97% of total reads **(MCF7 gDNA)**

We used BEDTools (Quinlan et al., 2010) and features of the genomic analysis of ChIP-Seq data (Euskirchen et al., 2007) for analysis of our data on DNA replication origin in the human genome. The reads of 42 bp listed in the table above were mapped to human genome build hg18 with Bowtie (Langmead et al, 2009; Langmead 2010). An input control was also sequenced and mapped to the genome to reduce the number of false positive results (This was not done previously for our pilot run on the Illumina GAIIx machine). The input control we have used is genomic DNA that has not been enriched for short nascent strands. Specifically, for MCF-7 genomic DNA (gDNA) we

used estrogen starved G0 cells and for the MCF-10A gDNA (described below in Task 2) we just used asynchronous cells. In brief, total gDNA was isolated from the cells, sheared by sonication to a size range of about 100-600 bp and the entire prep was taken through the Illumina library preparation. In the end, the gDNA library was size selected on 2% agarose for 200-500 bp fragments which were then sequenced.

Once treatment and control reads are mapped to the genome, statistically significant peaks, which are areas where treatment reads pile-up high over the control reads, are called with a peak caller. The 239 million mappable reads (see table above) were combined into a single file. This file was used to call peaks against the input reads with MACS. As MACS was designed for ChIP-seq experiments, we spent considerable time optimizing the parameters for Nascent Strand sequencing (NS-seq).

In year one, using only one lane of sequencing from the Illumina GAIIx which produced ~11.8 million mappable reads (see table above), we called 54,100 peaks (53,914 peaks after removing chrY peaks, which are artifacts in the MCF-7 context). Since we did not have an input control for that pilot experiment, some of the peaks were probably false positives. Moreover, we did not reach saturation, which required more sequencing. Finally, to be rigorous we needed to include sequence reads from biological replicates. To reach saturation and to include biological replicates, in year two of DOD support, three lanes on the Illumina Hi-Seq2000 platform were used to sequence 3 biological replicates of short nascent strand preparations from MCF-7 cells. The Hi-Seq2000 data gave 54.6 million, 84 million, and 89 million mappable reads for the three different lanes (see table above). Including our previous 11.9 million reads, this gave 239.6 million mappable reads. We also obtained ~180 million mappable input control reads (MCF-7 gDNA).

In the year one progress report, we presented our analysis of the 53,914 peaks from the pilot GAIIx sequencing, including median and mean peak widths and inter-origin distances. We also presented the data demonstrating good congruence between our reads and some known origins, including c-Myc, DBF4, DHFR, β-Globin, RPE, as well as Lamin B2, and Glucose-6-Phosphate Dehydrogenase. There have been a few other reports of mapping origins in the human genome but, surprisingly, we found that the overlap in origins was less than 40% at best between the datasets from the various labs (Cadoret et al., 2008; Karnani et al., 2010.; Mesner et al. 2011; Martin et al. 2011, Valenzuela et al. 2011; Mesner et al., 2013). Although some of these other reports only mapped origins to 1% of the HeLa genome (ENCODE project), the poor overlap of origins was especially surprising when comparing our data to that of Martin et al. (2011) who has also performed whole genome NS-Seq on MCF-7 cells. They used smaller nascent strands than us and we suspect that may have led to Okazaki fragment contamination in their samples. Indeed, their nascent strand enrichment was less than ours and not even reported in their paper nor were their results validated in their paper. We speculated that lack of saturation may be an explanation for poor congruence. However, now we believe that there may be some technical explanations (described below). Very recently, a paper appeared by Besnard et al. (2012) that mapped origins in the human genome. Remarkably, they found more than twice as many origins in the

human genome as discovered by us or Martin et al. (2011). This led us into experiments to try and unscramble the confusion in the field on the number and location of replication origins in the human genome. We now suspect problems arising from the technique of nascent strand DNA preparation and also from analysis of the sequencing data as the explanation for the poor congruence between origin datasets from the different groups. Our experiments and analysis on this issue are described below.

Technical issues in NS-Seq:

It has been reported that there is a base composition skew in metazoan replication origins and that they are GC-rich (Cayrou et al. 2011). Moreover, the origin G-rich repeated elements (OGREs) predict G-quadruplex (G4) structures at origins (Cayrou et al., 2012a). The recent paper by Besnard et al. (2012) claimed that a consensus G-quadruplex-forming DNA motif can predict the position of replication origins in the human genome. We suspected that this result from others might be a technical artifact. Lambda exonuclease (Lexo) is used in NS-Seq protocols to enrich nascent strands by virtue of their resistant to Lexo degradation because of the 5' RNA primer that protects them. If G-quadruplex DNA is also resistant to Lexo digestion, then G-quadruplex DNA will also be present in the nascent strand enriched sample after Lexo digestion, thus contaminating the sample. Postdoc Michael Foulk in our lab has done experiments to explore this possibility, and his data support the likely contamination of nascent strand DNA preparations by G-quadruplex DNA. The plasmid pFRT.Myc contains a G quadruplex on each strand (Pu27 and Pu30); Mike cut this plasmid with the restriction enzyme BglII to linearize it. Samples were run on a gel with or without Lexo digestion. Samples were boiled and placed on ice to favor intramolecular formation of the G-quadruplex and revealed products on the gel that would be predicted if Lexo stopped digesting the DNA when it came to the predicted position of a G-quadruplex roadblock (**Figure 1**). Mutation of the G-quadruplex sequences allows Lexo to digest the DNA, supporting the conclusion that G-quadruplex structure acts as a roadbock to impede Lexo digestion (**Figure 1**). Therefore Lexo digested DNA is enriched in G-quadruplexes in addition to newly synthesized DNA (**Figure 2**). As a result, when genomic DNA is obtained from non-replicating (G0 phase) MCF-7 cells and used for NS-Seq, there will be G-quadruplex peaks in addition to replication origin peaks (**Figure 3**). Moreover, when the sample of non-replicating DNA is digested with Lexo and subjected to NS-Seq, it will display the G-quadruplex peaks (including those that coincided with replication origins) but not origins themselves since the sample came from non-replicating cells (G0)(**Figure 4**). The peaks from the non-replicating cells whose genomic DNA was digested with Lexo can be used to correct the NS-Seq data set from replicating cells (**Figure 5**). When this correction is carried out for the well-studied c-Myc locus, we could identify which of the three peaks from the uncorrected data set remain in the corrected set. As shown in **Figure 6,** the peak on the left disappears as it was caused by Lexo stopping at G-quadruplexes that map to this area, the middle peak is diminished in size after correction for the Lexo bias and the peak on the right remains as it is a replication origin with no overlapping G-quadruplex. When applied genome-wide, the Lexo biases (inability to digest through GC-rich DNA and G-quadruplexes) can be seen from NS-Seq on non-replicating DNA digested with Lexo (**Figure 7** and **Figure 8**). Therefore, the nascent strand-independent lambda

exonuclease biases can explain the correlation of replication origins with G-quadruplexes reported by others (**Figure 9**). We observe that the NS-seq peaks of replication origins after our correction for Lexo biases are no longer GC rich (**Figure 10**).

Appendix 1 contains the NS-Seq data and analysis for the 3 replicate samples from Lexo-digested DNA from replicating MCF-7 cells. FACS analysis revealed that 30-40% of the cells were replicating. The contents of **Appendix 1** include:

To sum up, the correct way to combat the problem of Lexo biases for NS-seq is to sequence genomic (non-replicating) DNA that is enriched for G-quadruplex (by

boiling and quenching on ice followed by Lexo digestion) to map these structures in the MCF-7 genome to analyze if they are coincident or not with the replication origins already mapped from enriched nascent strand preparations and normalize the NS-Seq data from replicating cells accordingly. Another cautionary note is that Cayrou et al. (2011, 2012a and 2012b) boils her replicating DNA for 15 minutes before sucrose gradient centrifugation, and she uses pH 9.4 (rather than pH 8.8 recommended and used by us) for Lexo digestion. These harsh treatments could degrade RNA primers on the nascent DNA, making it even more likely that what she has sequenced is simply G-quadruplex DNA rather than nascent strands to identify replication origins. Similar concerns apply to the experiments of Martin et al. (2011) and Besnard et al. (2012).

In addition to correcting the NS-Seq data by normalization for peaks that are found after Lexo digestion of genomic DNA from non-replicating cells, we have developed an experimental way to minimize the Lexo biases in the NS-Seq sample from replicating cells. Specifically, the G-quadruplex stability depends on the cation (Shim et al., 2009) and it is less stable in Na+ than K+ (**Figure 11**). Therefore, simply changing the Lexo digestion conditions from glycine buffer with Na+ rather than K+ greatly reduces the Lexo G-quadruplex bias (**Figure 12**).

Since Lexo is widely used to isolate newly synthesized DNA for NS-Seq as well as for quantitative PCR of nascent strand abundance to map replication origins, our findings of the biases inherent in this enzyme and the importance to use Lexo digested controls from non-replicating cells as well as our observation that the biases can be minimized by use of Na+ rather than K+ in the Lexo digestion buffer represents a very important advance for the field. We are preparing for publication the information on the Lexo biases and ways to overcome them. Our observations call into question the NS-Seq data published thus far by others. We are now poised to map the true replication origins in the MCF-7 breast cancer genome and compare it to replication origins in the MCF-10A genome for normal breast cells. We have frozen cells and DNA samples for both genomes that are ready for the corrected Lexo digestion for NS-Seq.

Computational issues in NS-Seq analysis:

Most of the peak-caller software used for computational analysis of NS-Seq data was developed for analysis of ChIP-Seq data and has to be optimized for NS-Seq applications. In our case, we have used MACS and the optimization of the variables as done by graduate student John Urban in our lab is described below. The NS-Seq data of Martin et al. (2011) appears to have used a variant of MACS. In contrast, the NS-Seq data of Besnard et al. (2012) was analyzed with Sole-Search which they did not optimize and simply used the default parameters. John Urban in our lab has computed the effect of using MACS as compared to Sole-Search for analysis of our data, and the result is very striking, as described in Task 2 of this report.

Optimizing the variables in MACS

(A) Redundant Reads - John has compared the effect of keeping just 1 (K1) or 3 copies (K3) of redundant reads and discarding the remaining redundant reads (usually thought to be an artifact such as from PCR). As shown in **Appendix 2 Figure 1**, as

expected, somewhat more reads are kept in K3 than K1 for both the experimental sample of the combined nascent strand reads as well as for the nonreplicating genomic DNA input used to correct for background. The K1 and K3 options were used for further analyses and little differences were found between them.

(B) Normalizing number of reads between the treatment (NS-Seq) and the control (input) –

MACS will either scale the counts toward the larger file or to the smaller file. For the current dataset, the larger of the two files is always the treatment (NS-seq). This means that scaling to the large file adjusts the read counts in the input file and that scaling to the small file adjusts the read counts in the NS file. Henceforth, we will use the term "toLg" mean that the 'scale to the large file' option was used and we will use "toSm" mean that the 'scale to the small file' option was used. Another option for dealing with the disparity between the number of reads in the two files would be to adjust the number of reads in each file before submitting them to MACS such that the files have roughly the same number after MACS filters redundant reads out. John Urban has created a python script that calculates this adjustment and used it in a separate analysis where we called peaks using reads from only a single lane instead of combining all NS samples (discussed more below).

With just the two options discussed above, there are 4 possible sets of parameters. The parameter sets will be named such that it describes how many redundant reads were kept first, then whether the counts were scaled toward the large or small file: "K_to___". For example, keeping just 1 read (K1) and scaling toward the small file (toSm) will be called K1toSm.

(C) P value to call peaks - MACS uses the Poisson distribution to call peaks. Briefly, the Poisson distribution deals with the probability of X events occurring in a given time/space interval given that the average number of events that occur is λ (lambda). For MACS, it is the probability of seeing X reads in a genomic interval given that the average number of reads in that interval is λ (lambda). The p-value, which is used to determine if the read count is statistically significant, is just the probability of seeing greater than or equal to X reads in a genomic interval given that the average number of reads in that interval is λ (lambda). P-values closer and closer to 0 are considered more and more significant (they are less so as they approach 1). MACS allows the user to pick a p-value cutoff, C, where read counts with p-values higher than C (closer to 1) are not called as peaks whereas those with p-values less than or equal to C are called as peaks. Unless otherwise stated, we used the p-value cutoff of 0.00001.

(D) Dynamic genomic interval for the read count - Not only does MACS use the Poisson distribution, it uses a "dynamic lambda", meaning that it does not necessarily use the same average read count for all genomic intervals. This is in contrast to a 'static lambda' such as only using genomic average, λ(genome) = [number of reads/genome size]*interval_size. Instead, MACS looks for local biases in the genomic interval it is currently looking in. It does this by calculating two more λ values: λ(small local window)

and λ(large local window). The size of these two local window sizes can be tweaked. These parameters are called “slocal” and “llocal”. The default slocal is 1000 bp. This is approximately twice the size one might expect ChIP-seq peaks to be. We expected ~1500 bp peaks so kept ‘slocal’ fixed at 3000 bp while tweaking only the llocal option. **Appendix 2 Figure 2** shows a curve of the number of peaks called as a function of llocal. We wish to maximize the number of true peaks called. Note that all four conditions have a slight elevation in number of peaks at llocal = ~50 kb. This llocal value will continue to be most interesting in subsequent figures.

(E) False Discovery Rate to identify the true peaks – **Appendix 2 Figure 2** shows that there are 77,000-84,000 peaks in the MCF-7 genome from our combined NS-Seq data. How many of these are true peaks (bona fide replication origins) and how many are false positives? Using MACS, John Urban calculated the False Discovery Rate (FDR). As shown in **Appendix 2 Figure 3**, the expected number of true origins based on the data stayed somewhat constant for all 4 conditions (with llocal=~30 kb) with a range from ~66,000 to ~70,000. Note that all conditions had a slight elevation of true peaks and a slight dip in false peaks around llocal=50 kb. As will be seen below, taken together this means that there is also a slight decrease in FDR at this llocal value.

All 4 parameter sets have a slight dip in FDR at llocal = ~50 kb (**Appendix 2 Figure 4**). The “K1” sets have ~15% FDR after llocal = ~40-50 kb while the “K3” sets are slightly higher at ~16-17% FDR. When llocal = 50 kb, we have shown that the there is a slight elevation in the number of peaks, that the expected number of true positives is slightly elevated, that the expected number of false positives dips, and that as a result the FDR is most often lowest near llocal= 50 kb. This is why we chose to further explore the llocal = 50 kb sets for all conditions.

(F) Confirmation that the 50 kb llocal sets are appropriate for calling peaks. We wondered how different the llocal=50kb sets of a given condition (e.g. K1toLg) were from the other sets from the same condition when different llocal values were used. Note that for each condition the llocal=50 kb set contained the most peaks. Therefore, we asked whether or not the smaller sets were all proper subsets of the llocal=50 kb set – i.e. does the llocal=50 kb set contain every peak called from the smaller set in question? If not, we wanted to know how many peaks from the smaller set were not in the bigger set and whether that was lower than the expected number of false peaks in the smaller set. Each llocal value should have some unique peaks as a result of the differences in “dynamic lambdas” used. If the number of unique peaks exceeds the expected number of false peaks in a set, then that set would be considerably different from the llocal=50 kb set. This would be problematic because we would not have a way of knowing, which set was more representative of the truth. If the sets are reasonably similar, then choosing a set becomes more arbitrary and choosing the set that minimizes FDR while maximizing the number of peaks makes most sense. In such a case, we could move forward.

The peak sets are written out in BED files. Suffice it to say that each line in a BED file represents a peak and that the first three columns state the genomic

coordinates of that peak by specifying the chromosome, the start position, and the end position respectively. BEDtools is a bioinformatics program that can manipulate BED files in numerous ways (Quinlan et al., 2010). We used BEDtools to compare two sets of peaks at a time to see how many peaks were shared in common between the sets. In **Appendix 2 Figure 5** we report how many peaks in the smaller set were NOT in the larger llocal=50 kb set. The conclusion is that the 50 kb sets are appropriate to use in peal calling.

We next wanted to make sure that the peak sets did not vary much between conditions. If the peak sets were much different from each other between K1 and K3, then once again determining which of among the sets more reflected the truth would not be straightforward. However, if the sets did not vary considerably between conditions, then it simply would be arbitrary in choosing a set. One would just pick one that minimized FDR while maximizing peak calls. **Appendix 2 Figure 6** shows this analysis, which is analogous to Figure 5. This time the K3toLg set was biggest set so all smaller sets were compared to it. All sets were found to be reasonably similar. As the K1 sets had lower FDR, one of these was chosen as our final set. Scaling to small is supposed to have higher specificity and lower FDR. Nonetheless, we do not necessarily see this for the K1 sets. The K1toLg set actually seems to have a lower FDR.

(G) Number of peaks in each of these sets as a function of FDR%. This was analyzed primarily to access a given FDR set based on the present need. For example, one might prefer the 5% or 1% FDR sets when looking for a motif, but might prefer the entire set when calculating inter-origin distance. Moreover, we were able to see that the higher quality sets (e.g. 5% FDR) often vary much less between different conditions such as biological replicates (**Appendix 2 Figure 7**).

We also looked at how choosing different p-value cutoffs would affect number of peaks called as well as FDR (**Appendix 2 Figure 8**). This was done on the K3toLg set before we decided on the K1toLg. Nonetheless, the trend should be similar in all 4 conditions we considered, as they were all similar. Note that our p-value cutoff was 0.00001 and that log10(0.00001) = -5. In other words, the most stringent p-value cutoff is leftmost and the p-value cutoffs get more and more lax as it goes right along the x-axis to log10(0.1) = -1. What is interesting about what we see here is that the expected number of true peaks remains relatively constant while the expected number of false peaks grows with less stringency solely contributing to the rising total number of peaks. This seems to indicate that even more stringent p-value cutoffs might keep the same number of expected true positives while reducing further the expected number of false positives until the total number of peaks and expected number of true peaks are approximately the same. We are yet to do this in large part because another peak caller we use, Sole-Search (discussed later), calculates FDR in a different way and gives completely different FDR estimates (e.g. 0.001% instead of 15%) while the regions with peaks remain relatively the same. Moreover, **Appendix 2 Figure 9** of the expected numbers of false and true positives normalized by the total number of peaks shows FDR and “TDR” respectively as functions of pvalue cutoff. They seem to be leveling off in the left-direction indicating that the number of true peaks and total peaks will not

actually converge, at least not until after most true and false positives are eliminated from consideration.

Biological and technical variation

We analyzed the biological variation in three different samples of MCF-7 nascent strand DNA. Having set the variables as described above, for each of the separate samples we used MACS to call peaks from the mappable reads after correction for redundant reads and subtraction of background from nonreplicating genomic DNA (**Appendix 2 Figure 10**). The first DNA preparation (Rep1) was prepared by the BND-cellulose protocol, whereas the next two samples (Rep 2, Rep 3) were preparedsubsequently. Rep2 and Rep 3 turned out to be more similar to each other, as now described. We calculated how many peaks from a given set (row) were represented by another given set (column) (**Appendix 2 Figure 11**). This analysis is performed using BEDtools (Quinlan et al., 2010). This is followed by a table that instead shows the percent of the given set (row) that is represented in another set (column) (**Appendix 2 Figure 12**). The diagonal from top-to-bottom, left-to-right in the first table(Figure 12) is in bold because this shows the total number of peaks as the row set and column set are the same at these intersections. In the second table (Figure 13) this is evident as the diagonal has 100% in all boxes.

It is clear that Rep2 and Rep3 are more like each other than they are like Rep1. That Rep2 and Rep3 contribute ~173 million (173,135,267) mappable reads while Rep1 contributes just ~66.4 million (54,642,570 + ~11.8 million from GAIIx) speaks to why higher percentages of peaks from these files are represented in the Combined file and why higher percentages of peaks from the Combined file are represented in these files despite that they have less peaks than Rep1.

Next to look at the technical reproducibility, we compared the original GAIIx run reported last year with the HiSeq2000 Rep1 data treated in the same way as GAIIx. The HiSeq Rep1 data and the GAIIx data are derived from the same NS sample. We did not have input data for the GAIIx analysis. Therefore, in a separate analysis on the Rep1 data, we treated it as if there was no input control. Below we show the comparison to find how many peaks in the GAIIx set (54,100 peaks) are represented in the larger HiSeq2000 set (117,446 peaks):

Rep1	HiSeq200
GAIIx	51544
Percent	95.275416

This shows that 51,544 peaks out of the 54,100 peaks in the GAIIx set (~95.3%) are represented in the HiSeq set. Therefore, this high throughput sequencing of a NS sample is highly technically reproducible. Thus, sequencing adds a small amount of technical variability compared to the variability between sample preparations, which is a mixture of biological variability and that introduced by the procedure and/or by using different nascent DNA isolation procedures.

Saturation

We wanted to know whether or not we have reached saturation, where saturation is defined as reaching an area of diminishing returns with more sequencing. To do this, the reads in the Combined file were shuffled to mix reads from all experiments together. Otherwise, the reads from each experiment are just stacked on top of each other in the Combined file – e.g. first the GAIIx reads, then the HiSeq2000 Rep1 reads, etc. Why shuffle them first? To certain extents, the analysis of lower read counts and number of peaks they give rise to has already been done. Rep1 had ~43.5 million reads, that were mappable and passed MACS filtering, to call peaks while Rep2 and Rep3 had 65 and 69 million. Rep1, 2, and 3 gave rise to ~80,000, ~55,000, and ~63,000 peaks respectively while the Combined set of mappable reads that passed filtering (~171 million) had ~79,000 peaks, suggesting that we have reached saturation (**Appendix 2 Figure 13**). However, that the number of peaks seems to have leveled off also means that combining reads from biological replicates reduces the amount of spurious peaks called in a given replicate. For example, only ~65% of the peaks in Rep1 are represented in the Combined set (see analysis above). In other words, 35% of the peaks were not substantiated when biological replicate reads were added into the analysis. This is part of the power of combining reads from biological replicates. Therefore, it would be interesting to look at how many peaks are called when 10 million of the shuffled combined reads are used, 20 million, 30 million, and so on up to 230 million. Note that because (i) not all of the reads are mappable and (ii) there will be increasing rates of redundant read events with increasing number of reads, the true number of reads is equal to the number of mappable reads that pass the redundant reads filter in MACS set to keep only 1 read. **Appendix 2 Figure 14** shows peaks when no input control is used. This simply means all possible peaks (true and false positives) are included in count. We are in the midst of performing this saturation analysis using input reads as well. It appears as though we have come close to saturation. Whereas the first 80 million reads gave rise to ~90,000 peaks, the second 80 million reads only gave rise to ~104,000 or just 14,000 more. Another 80 million would give rise to even less additional peaks and most would be very low enrichments that became statistically significant due to the large increase in sample size (~240 million reads).

Task (2) Comparison of replication origin maps between breast cancer (ER+, ER-) and normal breast cells. These results would indicate if replication origin usage changes between normal and breast cancer cells, and if it varies between ER positive and ER negative breast cancer cells.

We are nearing completion of subtask (2a) to map replication origins in an ER+ breast cancer cell line --- namely MCF-7 (see Task (1)). Due to our discovery of the biases in Lexo digestion and the steps we had to take at the experimental and computational level to correct the NS-Seq data (see Task (1)), this has delayed the other tasks in this grant. The experiments in subtask (2b) to map replication origins in ER- breast cancer cells (e.g., MDAMB231 cells, SKBR3 cells) have been deferred since they are peripheral to the more important comparison in subtask (2c) to map replication origins in normal breast cells (MCF-10A). The similarities and differences in replication origin maps for the ER+ MCF-7 breast cancer cell genome compared to the origin map for the normal breast cell genome (MCF-10A) addresses whether replication origins

differ in different cell types, especially comparing breast cancer cells to normal breast cells.

MCF-10A

In year two we isolated nascent DNA as well as input genomic DNA from normal human breast cells (MCF-10A). We sequenced this material using the HiSeq2000 platform. The data obtained from this sample is presented in Appendix 2. However, we need to obtain more biological replicates and correct the data for the Lexo biases noted above. **Appendix 2 Figure 15** shows the results for the preliminary MCF-10A NS-Seq data.

Data analysis with Sole-Search vs MACS

As indicated earlier in this report, a recent paper by the Lemaitre group (Besnard et al., 2012) obtained more than double as many peaks as us or Martin et al. (2011). Where we were calling nearly 80,000 peaks for MCF7 and nearly 68,000 peaks for MCF-10A, they were calling 200,000 to 250,000 for their cell lines. Though they used different human cell lines than us, there is no reason to expect this massive difference. A notable difference is that they used Sole-Search (with default parameters) rather than MACS as the peak caller. Below we describe our calculations to compare the effect of using the MACS vs. Sole-Search peak caller. To do this, we used our data with Sole-Search.

Sole-search has fewer parameters than MACS to tweak. Lemaitre's group kept all default parameters:

```
Permutation:5
Fragment:200
AlphaValue:0.0010
FDR:0.0001
PeakMergeDistance:0
```

We did roughly the same for our MCF-7 data:

```
Permutation:5
Fragment:350
AlphaValue:0.0010
FDR:0.0001
PeakMergeDistance:0
```

The difference is in bold. We used 350 because it accurately reflects our average fragment length. This parameter only eliminates peaks < 350 bp in length (or < 200 bp in Lemaitre's case).

Another difference is that we provided our own MCF-7 input control reads while the Lemaitre group used the generic genomic reference reads provided by the software. These reads are sampled from input controls from an array of cell lines. Ultimately, this should cause some problems because it does not account for the specific biases (such as amplifications) in the genome that the nascent strands were purified from. This

difference will be explored further when we discuss the differences for our MCF-7 set when using our own input or using the generic input from the Sole-Search software. For now, we wanted to keep everything as similar to the MACS analysis as possible. This would tell us if the different methods of identifying peaks were responsible for the difference.

The results of using Sole-Search on our NS-Seq data for MCF-7 are striking:
MappedReads: 239566823
UniqueReads: 165908046
Number of peaks: 280,368
Average peak height: 52.88
Median peak height: 49
Highest peak: 229
Lowest peak: 23.2016210739615
Average peak width: 983.39

As shown in **Appendix 2 Figure 16**, there were over 280 thousand peaks when using Sole-Search with the same parameters to the Lemaitre group. Moreover, if we change the Sole-Search FDR parameter from 0.0001 to 0.001 (less stringent) or 0.00001 (more stringent) we get 334,197 and 258,243 peaks from our MCF-7 data respectively.

We also used Sole-Search on our MCF-10A data (**Appendix 2 Figure 17**). Specifically, for MCF-10A we used all of the same parameters as for MCF-7. As compared to the 67,812 peaks called by MACS, Sole-Search called 110,212 or 219,637 or 288,567 peaks for MCF-10A sets with FDR equal to 0.00001, 0.0001, and 0.001 respectively. Notice that the FDR reported for MACS was very high. Nonetheless, almost all of the MACS peaks are found within the Sole-Search set, as described below. The difference in numbers of peaks is mostly a consequence of having multiple smaller-width Sole-Search peaks in the same region as one larger-width peak from MACS (discussed in more detail below). Therefore, FDR between peak callers is not comparable. Each computes FDR differently and it is not apparent which approach is more appropriate or accurate.

It is seen for both cell lines that the different peak callers indeed give rise to different numbers of peaks when all of the data provided is the same. This is important because the number of peaks is used as a proxy for the number of origins of replication in the human genome. Moreover, the peak sets are used to estimate inter-origin distance, the average of which will be much smaller for the Sole-Search set, as well as for motif discovery and other downstream analyses. Finally, the peak locations are used to see what other genomic features they correlate with. These different outputs could lead to different conclusions – one output may give rise to a false correlation or break a true correlation.

We wanted to know if, despite the large difference in the number of peaks, if the peaks were in the same genomic regions. First, we tested how much peaks in the

MACS set overlapped with peaks in the corresponding Sole-Search set and vice versa. High percentages of overlapping peaks would mean that, despite the large difference in number of peaks, that peaks were being called in the same regions. An overlap is counted if a peak in one set, overlaps at least 1 peak in the other set by at least 1 bp. We found that almost all of the MACS peaks were found in the Sole-Search set, but only about 60% of the Sole-Search peaks were found in the MACS set for MCF-7 (**Appendix 2 Figure 18**) and for MCF-10A (**Appendix 2 Figure 19**).

We next looked to see if each of the peak-callers called peaks in regions known to have replication initiation activity. If the sets from the 2 different peak callers are not similar as determined by the overlap test described above, then which might be more accurate (as determined by having peaks at know origin sites)? Conversely, if they are determined to be similar, do they both have peaks in these regions of known origins? To determine whether or not one, none, or both had peaks in regions known to have origin activity, we visualized the peaks from the MCF-7 and MCF-10A sets for both MACS and Sole-Search peak callers in the IGV browser and looked at 3 specific sites: the c-Myc locus, the HBB (beta globin) locus, and the RPE locus (**Appendix 2 Figures 20-22**, respectively). Both peak callers had peaks in all 3 of these places in both cell lines. This also visually shows that Sole-Search places many smaller-width peaks inside single, wider peaks called by MACS. It is clear that MACS has poorer resolution, but though the Sole-Search resolution seems to have finer resolution, it is unclear whether it parsed up the region too much, particularly at the c-Myc locus where some of the peaks called by Sole-Search do not coincide with experimental data. The top-most row always displays RefSeq genes (blue) at the given locus. It is then followed by MACS peaks (red) for MCF-7, Sole-Search peaks (green) for MCF-7, MACS peaks (red) for MCF10-A, and Sole-Search peaks (green) for MCF10-A, respectively. Above the rows showing genomic features (genes and NS peaks) is a representation of the chromosome and the width of the locus being viewed.

Finally, we looked at the density of peaks along the genome to see if the density rises and falls with each other. In other words, do the density curves of peaks along the genome visually correlate with each other? **Appendix 2 Figure 23** shows the densities of peaks along the genome starting with chromosome 1 and going up through chromosome X. The density of RefSeq genes is first shown in blue. Next, in red, is the density of the MACS MCF-7 peaks followed by the density of MACS MCF-7 peaks when they are randomly shuffled across the genome. In green, the Sole-Search peaks for MCF-7 are shown followed by randomly shuffling them across the genome. Notice that the nascent strand peaks for both peak callers have similar profiles and that this profile is similar to the density of RefSeq genes. Moreover, this similarity is broken if the peaks are randomly shuffled. This suggests once again that the two different peak callers are detecting nascent strand signal from the same regions of the genome though they are parsing up these regions slightly differently leading to different numbers of peaks.

Visually, it appears that the peak densities of both peak callers go up and down with each other. For a more quantitative description of how the densities go up and

down with each other, we tested for correlation using two tests: Pearson's r and Spearman's rho. The Pearson test for correlation assumes a linear relationship whereas the Spearman test does not. Instead, the Spearman ranks the scores in each set and then tests if the ranks of the different sets go up and down with each other. Both tests were used to test for correlation and found a moderately strong positive correlation between the peak densities of both peak callers (**Appendix 2 Figure 24**). Note that for zero peaks called by MACS, several can be called by Sole-Search, helping to explain the 40% additional peaks called by Sole-Search as compared to MACS (Supporting Data Figure 19). If one of the peak sets is randomly shuffled across the genome, the correlation is broken (**Appendix 2 Figure 25**).

The effects of using the 'generic input' option in Sole-Search on our data

As mentioned, the Lemaitre group (Besnard et al., 2012) did not use their own input control sequencing reads. Instead, they specified to Sole-Search to use generic reads that the creators of Sole-Search have amassed from various cell lines. The question becomes, "Does using the generic input option change our results in comparison to using our specific set of control reads when all other parameters are kept the same?" Sole-search called 280,368 peaks when our own specific input control reads for MCF-7 were provided, but only 194,815 peaks (~69.5% the size) when their generic cell line reads were used as a control (**Appendix 2 Figure 26a**). This means that when using generic input, Sole-Search called 85,553 less peaks than when using the specific input, which indicates it has less sensitivity when generic input is used. The size of the peak set called with generic input is only ~69.5% the size of the specific input peak set (**Appendix 2 Figure 26b**), but what percent of the specific input set is actually covered by the generic input set? The answer is only ~57% (**Appendix 2 Figure 27**). What percent of the generic set is represented in the specific set? The answer is ~91.1% (Supporting Data Figure 28). Taken together, this shows that ~91.1% of the peaks called in the generic set agree with peaks in the specific set covering just 57% of the specific set. Relative to the specific set this implies that using the generic input leads to Sole-Search having a false negative rate of ~43% (i.e. the probability of not calling a peak that would be called with specific input is ~43%) (**Appendix 2 Figure 28**). Moreover, that 91% of the generic set is represented in the specific set leaves ~8.9% unique to the generic set. Relative to the specific set, this means ~8.9% of the peaks are false peaks and that the minimum FDR, defined as the number of false peaks divided by the total number of peaks in the set (false/total), is ~8.9%. However, the FDR may be even higher. 91% of the generic set overlap 57% of the specific set. This means that there are instances when more than one generic peak overlaps the same specific peak. If one takes the number of peaks represented in the specific set and divides by the total number in the generic set, it gives an approximation to the percent of non-redundant peaks in the generic set that overlap a peak in the specific set, just 82%. That means the FDR relative to the specific set could be as high as 18%. Taken together, all of this could imply that even though the Lemaitre group report the FDR given by Sole-Search (0.0001), it may be that the FDR is far higher. This is in addition to having a false negative rate relative to the specific set that implies they could not have reached saturation of origins – just saturation at their level of sensitivity.

Nucleotide composition of the peaks

To begin analyzing the sequences of our NS peaks for both MACS and Sole-Search, we aligned all of the peaks by either aligning the peak summits (nucleotide of highest coverage) and/or the peak centers (position in middle of peak length). From the given alignment focal point, we looked at the first 2000 nucleotides (nt) in both directions. Therefore, we looked at the 4000 nt centered around the summit/center for all peaks to find the nucleotide proportions at each position. This provides insight into whether or not there are any skews in the nt distribution in the peaks deviating from the random background distribution. To show the random background distribution, the peaks and/or peak summits were also randomly shuffled around the genome and treated the same as discussed above.

When this analysis was done for MCF-7, the following distributions were seen:

(1) MACS MCF7 NS peak summits (**Appendix 2 Figure 29**)
(2) MACS MCF7 NS peak centers (**Appendix 2 Figure 30**)
(3) MACS MCF7 shuffled NS peak summits (**Appendix 2 Figure 31**)
(4) Sole-Search MCF7 NS peak centers (**Appendix 2 Figure 32**)
(5) Sole-Search MCF7 shuffled NS peak centers (**Appendix 2 Figure 33**)

From the random distribution obtained from shuffling peaks, it is clear that the background proportions for A and T are ~29% each and for G and C are ~21% each regardless of position. In other words, at random, the nt distribution is position-independent. However, when centered at summits or peak centers, the distribution becomes position-dependent. It is clear that within 500-1000 nt from the peak center in both directions there is a non-random nt distribution that increases in GC content as it approaches the center.

When this analysis was done for MCF-10A, the following distributions were seen:

(1) MACS MCF-10A NS peak summits (**Appendix 2 Figure 34**)
(2) MACS MCF-10A peak centers (**Appendix 2 Figure 35**)
(3) Sole-Search MCF-10A peak centers (**Appendix 2 Figure 36**)

The background distribution stays the same as shown above. The same trend of higher GC content as one approaches the summit/center is also seen in the MCF-10A data.

To explore whether the nt skew found in the summit/center of peaks was meaningful or whether it was an artifact, we also looked at peaks that were not derived from Nascent Strand enrichment, using genomic DNA (gDNA) as a control. We used MACS to call peaks from just the control input reads. We did this originally to approximate areas of amplifications. The peaks were called using the static lambda (genome background average read count for a given interval). A dynamic lambda would not work to call amplifications because it would use the local average read count to see whether there was significant enrichment over background in a given area. However, an amplicon is not significantly enriched over itself. It is significantly enriched over the genomic average. These amplicon peaks were assessed the same as described above for their nt composition centered around the summits (nt of highest coverage). It is clear, that the same GC bias arises around the summit of these peaks, which have not come from our NS isolation protocol. The results shown in **Appendix 2 Figures 37 and 38**

argue that the task of centering peaks around their summits/centers in combination with biases common to all Illumina preparations (PCR bias, sequencing bias) at least partially explains this distribution. In other words, it argues that this GC skew by summits and centers that has been reported by others is an artifact.

As another control for the GC skew, we looked at the "negative peaks" MACS called for our MCF-7 and MCF-10A datasets. Negative peaks arise from the FDR approximation process of MACS. First, it calls all peaks in the treatment file using the input file as background. Regions in the treatment file enriched over the same regions in the input file are called as peaks. To approximate how many of the peaks called are 'false peaks', areas enriched over background by chance, MACS swaps the roles of the treatment and input files. Now it looks for regions in the input file that are significantly enriched over the same regions in the treatment file using all of the same parameters. MACS considers these to be false peaks by definition. It then uses this number as a proxy for the number of false peaks that were called in the treatment file to estimate the FDR of the entire set:
FDR = #false/#total. It also provides all of the locations for these "negative peaks". Negative peaks are interesting for this analysis because they arise in regions that are seriously deprived or depleted of reads from the Nascent Strand sequencing relative to the number of reads at those regions in the input control. This argues that these are regions that lack replication origins. Moreover, the negative peak summits in these regions may arise at non-random locations as reads are non-uniformly distributed across the genome, most likely due to PCR or sequencing biases. When the negative peaks are centered at the summits, we see similar nucleotide skews away from the random distribution for both MCF-7 (**Appendix 2 Figures 39 and 40**) and for MCF-10A (**Appendix 2 Figures 41 and 42**). This argues that the center/summits of peak regions are, in general, non-randomly enriched at areas within those regions of higher GC content. Therefore, though our peaks may indeed localize at or near origins of replication, where the peak summits and/or centers is not necessarily meaningful. Aligning the peaks by either gives rise to an artifact. This means another approach needs to be taken to align the peaks such as a Hidden Markov Model approach that will align the peaks by state paths modeling nt compositions or a motif.

We next examined our MACS MCF7 set for any base composition bias for the "best" peaks as determined by p-value scores (**Appendix 2 Figures 43 and 44**). Interestingly, they show slightly different distribution patterns than all the peaks combined. However, both still have a GC rise in the middle. The top 114 peaks show a lot of variation from one position to the next, but have a general trend of AT content that is higher than the genomic average. That is interesting as origins of replication from all bacteria, bacteriophages, animal viruses, and unicellular eukaryotes studied thus far have high AT content, presumably to facilitate unwinding of the DNA.

Task (3) Correlation of origin map data with sites of (a) DNA amplification and (b) estrogen receptor binding. These data will support or refute the hypothesis that ER may bind next to the replication machinery and induce DNA amplification. This analysis was originally scheduled for year two. However, because of the unanticipated delay

necessitated by our forging new ground to refine technical and computational issues for NS-Seq, task (3) is deferred until the NS-Seq data after Lexo bias correction is completed. We will compare the origin map data to data that already exists on sites of DNA amplification (to identify amplification origins) as well as confirm and expand these data using our own data on the number of reads from sequencing bulk genomic DNA from the various cell lines we are using. This information will, in turn, be compared to existing data on sites of ER binding (Lin et al., 2007; Welboren et al., 2009). It may prove necessary to undertake some ChIP (chromatin immunoprecipitation) experiments for validation of ER binding, though not proposed in the original grant application. These data will indicate if a correlation exists between ER binding and origins that re-replicate (amplify), thereby testing our hypothesis. We are submitting a letter of intent for a grant application to DOD BCRP to fund the studies needed to complete these studies.

As indicated in the application for the present grant, we have begun to stockpile surgically derived breast cancer tissue (**Table 1** in Supporting Data), provided to us a residual, de-identified tissue from surgeons Theresa Graves and Maureen Chung and pathologist Shamlal Mangray, all from Rhode Island Hospital which is affiliated with the Brown University Medical School. In the future, beyond the scope of the present grant, we will use samples of this tissue to refine the methodology we developed in task (1) for use on surgical specimens. Pilot runs will be initiated in ER + human breast cancer tissue to map replication origins and sites of DNA amplification to compare to matched normal breast tissue from the same patient. These data will be expanded in future studies to reveal if novel origins are used for re-replication and if they correlate with ER binding sites adjacent to them. This information will have clinical importance.

Concluding remarks – The finding that the transcription factor c-Myc interacts with the pre-replication complex to control DNA replication (Dominguez-Sola et al., 2007; Lebofsky and Walter, 2007) and that the androgen receptor interacts with MCM7 of the pre-replication complex (Shi, 2008) provides precedence for our hypothesis that the ligand-bound estrogen receptor may play a direct role in regulating replication origins beyond its traditional role as a transcription factor. We are grateful for the DOD funding that allowed us to initiate experiments to test our hypothesis and hope that future DOD funding will allow us to complete these studies.

KEY RESEARCH ACCOMPLISHMENTS

- Prepared samples of nascent strands for three biological and technical replicates for replicating DNA from MCF-7 breast cancer cells and one sample from MCF-10A normal breast cells.

- Used the three MCF-7 samples and one MCF-10A sample for NS-Seq on the Illumina HiSeq 2000 machine.

- Validation of the NS-Seq results by finding known replication origins in our data set - We have validated NS-Seq on known origins, including Myc, DBF4, DHFR, β-Globin, RPE, as well as Lamin B2, and Glucose-6-Phosphate Dehydrogenase.

- Compared computational approaches to analyze NS-Seq data from these samples and concluded that MACs is preferable over Sole Search.

- Comparison of our data to the data sets of other labs to map replication origins in the human genome. The data sets from Cadoret et al., 2008; Karnani et al., 2010., and Mesner et al. 2011 were based on using ENCODE (1% of the human genome) for HeLa cells, so finding only a small amount of overlap could be due to their use of a different cell line than that used by our lab. Moreover, even when comparing the results between these three data sets, there was not complete agreement, suggesting lack of saturation of the data. The Martin et al. (2011) data set used MCF-7 cells and was for the full genome, but did not give full overlap with our data. They did not show any data for validation of their results, and we suspect that they had contamination from Okazaki fragments as they selected small nascent strand DNA. We have spent considerable effort to analyze the effects of different peak callers (MACS and Sole-Search), and our results (Appendix 2) demonstrate that the larger number of peaks called by the Lemaitre group (Besnard et al., 2012) reflects their use of the Sole-Search peak caller.

- Analysis of base composition at the peak summits or centers. Our computational analysis reveals that the apparent GC skew at the peaks seen by other groups appears to be a computational artifact. In fact, the “best” peaks in our data set have do not have the GC skew, which is more consistent with the base composition of replication origins from bacteria, viruses and yeast.

- Performed experiments with a plasmid containing G-quadruplexes in the c-Myc locus and demonstrated that Lexo cannot digest through G-quadruplexes.

- Sequence analysis of genomic DNA from non-replicating MCF-7 cells (G0 phase) that was digested with Lexo revealed the biases inherent in this enzyme that has difficulty digesting through G-quadruplexes and GC rich DNA.

- Developed computational approaches to use the Lexo treated G0 DNA to correct the NS-Seq data from replicating MCF-7 cells.

- Carried out experiments that demonstrated reduction in the Lexo biases if the Lexo digestion buffer contains Na+ rather than K+.

- Further refinement in the method to isolate nascent (newly replicated) DNA to reduce Lexo biases and use a lowered pH to prevent RNA degradation during lambda exonuclease digestion.

REPORTABLE OUTCOMES

A paper is being prepared for publication to report the following:

- Demonstration that Lexo has an inherent bias and cannot digest through G-quadruplexes.
- Methods to combat this problem:
  (a) computational approaches to use the Lexo treated G0 DNA to correct the NS-Seq data from replicating MCF-7 cells.
  (b) Change in Lexo digestion condition to use a buffer containing Na+ rather than K+ to minimize the Lexo biases.

A methods review may be submitted for publication, describing

- Refinement of the method to isolate nascent (newly replicated) DNA using Lexo digestion and the appropriate Lexo treated G0 DNA as a control
- Discussion of the various computational approaches to analyze NS-Seq data.

A research paper can be submitted after the data and analysis is completed to report on

- Origin mapping genome-wide in MCF-7 breast cancer cells.
- Origin mapping genome-wide in MCF-10A normal breast cells.
- Comparison in the replication origin maps between the two different cell lines, allowing the comparison of replication origin usage between cancer and normal breast cells.

A research paper can be submitted after the data and analysis is completed to report on

- Genome-wide map of origins that map to regions of DNA amplification in MCF-7 breast cancer cells
- Comparison of these origins to the same origins from non-amplified regions in the MCF-10A normal breast cell genome to identify any sequence motifs that may lead to DNA amplification.
- Analysis of estrogen receptor binding sites in the MCF-7 breast cancer genome to see if there is a significant correlation with their juxtaposition to replication origins in regions of DNA amplification.

Completion of the last two studies will require additional funding.

Several presentations have been made at scientific meetings of our DOD-funded research:

(1) S.A. Gerbi, M. Foulk, A. Brodsky and B. Raphael (2011). Origins of DNA replication and amplification in the breast cancer genome. Department of Defense Era of Hope Breast Cancer Meeting (Orlando, FL), Poster P48-8.

(2) J. Urban, M. Foulk, C. Casella and S.A. Gerbi (2011). Mapping DNA replication origins to the human genome. Cold Spring Harbor Laboratory Meeting on Eukaryotic DNA Replication and Genome Maintenance, Poster p. 222a.

(3) S.A Gerbi, J. Urban and M. Foulk (2013). Mapping DNA replication origins in the human genome. American Society for Biochemistry and Molecular Biology annual

meeting; Experimental Biology 2013 meeting (Boston, MA), Invited talk: abstract 759.1.

(4) M.S. Foulk, J.M. Urban, C. Casella and S.A. Gerbi (2013). Mapping DNA replication origins in the human genome. Cold Spring Harbor Laboratory meeting on Eukaryotic DNA Replication and Genome Maintenance (September 9-13, 2013), Invited talk: p. 3.

The abstracts for each of these four presentations are below:

(1) S.A. Gerbi, M. Foulk, A. Brodsky and B. Raphael (2011). Origins of DNA replication and amplification in the breast cancer genome. Department of Defense Era of Hope Breast Cancer Meeting (Orlando, FL), Poster P48-8.

The fidelity of DNA replication is of paramount importance for normal function of a cell. Disregulation of replication can lead to DNA amplification that is a hallmark of cancer. When oncogenes are amplified, they promote growth of the cancerous cell. Hence, it is important to understand the mechanism underlying DNA amplification. We suggest that re-firing of an origin of replication may be an initiating event in DNA amplification. Our previous research on developmentally regulated DNA amplification in a model system of the fly *Sciara* demonstrated that a steroid hormone triggers re-firing of a DNA replication origin, resulting in DNA amplification. The steroid hormone estrogen has been implicated in breast cancer progression. Can our previous results in the fly serve as a paradigm --- can estrogen induce DNA amplification in breast cancer? We want to learn whether binding sites for the estrogen receptor are located adjacent to origins of DNA amplification in the genome of MCF7 breast cancer cells. Sites of DNA amplification and sites of binding of the estrogen receptor have already been identified in the MCF7 genome. To map the origins of DNA amplification requires that we map all replication origins in the MCF7 breast cancer genome. In order to identify origins of replication, we have made preparations of short nascent strands that will be sequenced using next generation sequencing technology. In brief, nascent DNA is resistant to lambda-exonuclease digestion because of the presence of a 5' RNA primer, allowing the parental DNA to be digested while the nascent DNA is untouched. In a preliminary experiment we were able to enrich (up to 19-fold) for nascent strands from asynchronously growing MCF7 cells that were subsequently sequenced by Illumina. Our data overlapped with 78 of the 283 replication origins identified in HeLa cells by Cadoret et al. (2008) in the ENCODE region of the human genome. Encouraged by these results, we optimized the protocol to further enrich for nascent DNA, adding precautions to stabilize the RNA primer on the nascent DNA. Using the c-Myc origin to assess for enrichment, we have produced several preparations with substantial enrichment (up to 100-fold). We have sent these nascent strands for sequencing using Helicos single molecule sequencing and are currently analyzing the data. We also intend to use Illumina to sequence nascent strands in the near future and compare the results between the two platforms. Mapping all the replication origins in the MCF7 genome will allow us to identify which origins occur at regions of DNA amplification and whether they reside in close proximity to estrogen receptor binding sites.

(Supported by DOD CDMRP log # BC097936)

(2) J. Urban, M. Foulk, C. Casella and S.A. Gerbi (2011). Mapping DNA replication origins to the human genome. Cold Spring Harbor Laboratory Meeting on Eukaryotic DNA Replication and Genome Maintenance, Poster p. 222a.

We have mapped replication origins in the human genome using next generation sequencing technology. Asynchronous MCF7 human breast cancer cells in log phase were used for preparations of short nascent strands for sequencing on the Illumina platform. Our nascent strand-seq ("NS-Seq") protocol is based on our earlier report (AK Bielinsky & SA Gerbi. 1998. Science 279:95-8) that nascent DNA is resistant to lambda-exonuclease digestion because of the presence of a 5' RNA primer. This allows the parental DNA to be digested while the nascent DNA is untouched. The nascent strands were size selected on gels for 1-2 kb, which gave greater origin enrichment than a 0.5-1 kb fraction that may have Okazaki fragment contamination. Using the c-Myc origin to assess for enrichment, the average of the several preparations used for sequencing had 54-fold enrichment of nascent strands. Interestingly, in assessing this enrichment at the c-Myc origin, we discovered that the preferred origin resided in the second exon of the gene while it was previously determined to reside in the promoter of the gene (in HeLa cells: L Tao et al. 2000. J. Cell Biochem 78:442-57). This observation was confirmed in our NS-Seq data, suggesting plasticity of origin usage at the c-Myc gene in different cell types. We identified 53,914 origins in the MCF7 genome, with a median width of 1.5 kb. Many known replication origins were present in our data set including c-Myc, DHFR, Dbf4, Lamin B2, beta-Globin and Glucose-6-Phosphate Dehydrogenase. There are varying degrees of overlap between our dataset and those of others (JC Cadoret. 2008. PNAS 105:15837-42; N Karnani et al. 2010. Mol Biol Cell 21:393-404; Mesner et al. 2011. Genome Res 21:377-89; MM Martin et al. 2011. Genome Res) as will be discussed.
(Supported by DOD CDMRP log # BC097936)

(3) S.A Gerbi, J. Urban and M. Foulk (2013). Mapping DNA replication origins in the human genome. American Society for Biochemistry and Molecular Biology annual meeting; Experimental Biology 2013 meeting (Boston, MA), Invited talk: abstract 759.1.

We will present our results using nascent strand sequencing (NS-Seq) to map DNA replication origins in the human genome. Nascent DNA strands were isolated from MCF7 and MCF10A cells to allow comparisons between breast cancer and normal genomes. NS-Seq employs lambda exonuclease (Lexo) to destroy parental DNA; newly replicating DNA has an RNA primer at its 5' end, rendering it resistant to Lexo digestion. We will compare our results to those of others and discuss the methodological and computational factors leading to the large discrepancy between their data sets. Our goal is to resolve the confusion in the field on the number and identity of replication origins. The fractionated size of the nascent strand sample is important to avoid contamination from Okazaki fragments. Also, non-replicating input DNA used as a control is important. Variables in the computational analysis will be discussed. We have analyzed whether DNA replication origins display a base composition bias, the average spacing between

replication origins, how many origins map to intergenic regions as compared to within genes and association of origins with the estrogen receptor.
(Supported by DOD Breast Cancer Program W81XWH-10-1-0463)

4) M.S. Foulk, J.M. Urban, C. Casella and S.A. Gerbi (2013). Mapping DNA replication origins in the human genome. Cold Spring Harbor Laboratory meeting on Eukaryotic DNA Replication and Genome Maintenance (September 9-13, 2013), Invited talk: p. 3.

We will present our new approach to and results from our nascent strand sequencing (NS-seq) experiments to map DNA replication origins in the human genome utilizing the power of the Illumina HiSeq2000 platform. As we described previously (Gerbi and Bielinsky (1997) Methods 13(3): 271-280; Bielinsky and Gerbi (1998) Science 279: 95-98), nascent strands are enriched by digestion of genomic DNA with lambda exonuclease (Lexo) to destroy parental DNA. In contrast, newly replicating DNA has an RNA primer at its 5' end, rendering it resistant to Lexo digestion. Initially, Lexo was used to map known origins at the nucleotide level of resolution. More recently, it has been used by several groups for genome-wide origin discovery. Of great significance to many approaches that use Lexo to map origins *de novo*, we now describe its nascent strand independent biases that give rise to many false positives in NS-seq and ways to overcome them. We demonstrate that Lexo stalls at G4 quadruplexes, which can readily form in the nascent strand preparation conditions. Thus, in addition to nascent strands, Lexo digestion also enriches fragments of DNA that have a G4 quadruplex near their 5' ends. Therefore, the Lexo-treated DNA preparations of nascent strands are contaminated with G4 quadruplexes as well as other sequences that are inefficiently digested by Lexo. We show that the adjustment of buffer conditions can minimize the stalling of Lexo at G4 structures. In addition, we present a computational approach to correct for Lexo associated artifacts. Specifically, we performed a genome-wide analysis to call peaks from DNA from non-replicating cells (G0) that has been subjected to the same steps of Lexo treatment as the experimental sample of nascent strands from asynchronous cells. We show that these peaks called in the absence of nascent strands can alone explain the correlation with predicted G4 motifs seen by other groups. Therefore, the coincidence between replication origins and G4 quadruplexes as reported by others is not supported by our data. With the corrections that we present here, a purer set of replication origins can be mapped in the genome. Our findings underscore the importance of using orthogonal approaches (i.e., assays not involving Lexo, such as nascent strand extrusion, bubble trap, and pull down of ORC-associated DNA) to validate origins newly discovered in the genome.

(Supported by DOD Breast Cancer Program grant W81XWH-10-1-0463 (to SAG) and NSF GRFP 201111-7803 (to JMU)).

CONCLUSION

We are forging new ground to refine the method for nascent strand sequencing (NS-Seq) and to optimize the experimental and computational approaches. Our data has revealed serious problems with Lexo biases that flaw the results published so far by other groups to map replication origins in the human genome. We have developed ways to overcome these problems. Therefore, when completed, our study will be the first corrected data set mapping replication origins in the human genome – notably in the MCF-7 breast cancer genome. These results will serve as the framework to test the hypothesis if there is a correlation between amplification origins and sites of estrogen receptor binding.

Personnel Paid From This Grant

DOD W81XWH-10-1-0463 (9/1/10-8/31/13)

PI and co-PIs:

Susan Gerbi	Professor of Biology (PI)
Alexander Brodsky	Assistant Professor of Medical Science (co-PI)
Benjamin Raphael	Associate Professor of Computer Science (co-PI)

Lab Personnel:

Yutaka Yamamoto	Research Associate
Michael Foulk	Research Associate
Jacob Bliss	Research Assistant
Souriya Vang	Research Assistant

REFERENCES

Besnard E., Babled A., Lapasset L., Milhavet O., Parrienello H., Dantec C., Marin J.M. and Lemaitre J.M. (2012). Unravelling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. Nature Struct. Mol. Biol. 19: 837-844.

Bielinsky A.K. and Gerbi S.A. (1998). Discrete start sites for DNA synthesis in the yeast ARS1 origin. Science 279:95-98.

Brewer B.J., Payen C., Raghuraman M.K. and Dunham M.J. (2011). Origin-dependent inverted-repeat amplification: a replication-based model for generating palindromic amplicons. PLoS Genet. 7(3):e1002016.

Cadoret J.C., Meisch F., Hassan-Zadeh V., Luyten I., Guillet C., Duret L., Quesneville H. and Prioleau M.N. (2008). Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proc. Nat. Acad. Sci. 105, 15837-15842.

Cayrou C., Coulombe P., Vigneron A., Stanojcic S., Ganier O., Peiffer I., Rivals E., Puy A., Laurent-Chabalier S., Desprat R. and M Mechali M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. Genome Res. 21: 1438-1449.

Cayrou C., Coulombe P., Puy A., Rialle S., Kaplan N., Segal E. and Mechali M. (2012a). New insights into replication origin characteristics in metazoans. Cell Cycle 11: 658-667.

Cayrou C., Grégoire D., Coulombe P., Danis E. and Méchali M. (2012b). Genome-scale identification of active DNA replication origins. Methods 57: 158-164.

Debatisse M. and Malfoy, B. (2005). Gene amplification mechanisms. Adv. Exp. Med. Biol. 570: 343-361.

Dominguez-Sola D., Ying C.Y., Grandori C., Ruggiero L., Chen B., Li M., Galloway D.A., Gu W., Gautier J. and Dalla-Favera R. (2007). Non-transcriptional control of DNA replication by c-Myc. Nature 448: 445-451.

Dorn E.S., Chastain P.D. 2nd, Hall J.R. and Cook J.G. (2009). Analysis of re-replication from redregulated origin licensing by DNA fiber spreading. Nucleic Acids Res. 37:60-69.

Dutta A. (2007). Chaotic license for genetic instability and cancer. Nature Genet. 39: 10-11.

Euskirchen G.M., Rozowsky J.S., Wei C.L., Lee W.H., Zhang Z.D., Hartman S., Emanuelsson O., Stolc V., Weissman S., Gerstein M.B., Ruan Y., Snyder M. (2007)

Mapping of transcription factor binding regions in mammalian cells by ChIP: comparison of array- and sequencing-based technologies. Genome Res. 17: 898-909.

Finn K.J. and Li, J.J. (2013). Single-stranded annealing induced by re-initiation of replication origins provides a novel and efficient mecvhanism for generating copy number expansion via non-allelic homologous recombination. PLoS Genetics 9(1): e1003192.

Gerbi S.A. and Bielinsky A-K. (1997). Replication initiation point mapping. Methods. 13, 271-280.

Gomez M. (2008). Controlled rereplication at DNA replication origins. Cell Cycle 7: 1313-1314.

Gomez M. and Antequera F. (2008). Overreplication of short DNA regions during S phase in human cells. Genes Dev. 22: 375-385.

Gopalakrishnan V., Simancek P., Houchens C., Snaith H.A., Frattini M.G., Sazer S. and Kelly T.J. (2001). Redundant control of rereplication in fission yeast. Proc. Nat. Acad. Sci 98: 13114-13119.

Green B.M. and Li J.J. (2005). Loss of rereplication control in Saccharomyces cerevisiae results in extensive DNA damage. Mol. Biol. Cell 16: 421-432.

Green B.M., Morreale R.J., Ozaydin B., Derisi J.L. and Li J.J. (2006). Genome-wide mapping of DNA synthesis in Saccharomyces cerevisiae reveals that mechanisms preventing reinitiation of DNA replication are not redundant. Mol. Biol. Cell 17: 2401-2414.

Green B.M., Finn K.J. and Li J.J. (2010). Lodd of DNA replication control is a potent inducer of gene amplification. Science 329: 943-946.

Hampton O.A., Den Hollander P., Miller C.A. et al. (2008). A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. Genome Res. (Epub. Dec. 3, 2008).

Hook S.S., Lin J.J. and Dutta A. (2007). Mechanisms to control rereplication and implications for cancer. Curr. Opin. Cell Biol. 19: 663-671.

Karnani N., Taylor C., Malhotra A. and Dutta A. (2007). Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas. Genome Res. 17: 685-676.

Karnani N., Taylor C.M., Malhotra A., Dutta A. (2010) Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. Mol Biol Cell. 21: 393-404.

Langmead B., Trapnell C., Pop M. and Salzberg S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol.10(3):R25.

Langmead B. (2010). Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics. Chapter 11:Unit 11.7

Lebofsky R. and Walter J.C. (2007). New Myc-anisms for DNA replication and tumorigenesis? Cancer Cell 12: 102-103.

Lin C.Y., Vega V.B. et al. (2007). Whole-genome catography of estrogen receptor alpha binding sites. PLoS Genet. 3(6): e87.

Liontos M., Koutsami M., Sideridou M., Evangelou K., Kletsas D. et al. (2007). Deregulated overexpression of hCdt1 and hCdc6 promotes malignant behavior. Cancer Res. 67: 10899-10909.

Lucas I., Palakodeti A., Jiang Y., Young D.J., Jiang N., Fernald A.A. and LeBeau M.M. (2007). High-throughput mapping of origins of replication in human cells. EMBO Rep. 8: 770-777.

Martin M.M., Ryan M., Kim R., Zakas A.L., Fu H., Lin C.M., Reinhold W.C., Davis S.R., Bilke S., Liu H., Doroshow J.H., Reimers M.A., Valenzuela M.S., Pommier Y., Meltzer P.S. and Aladjem M.I. (2011). Genome-wide depletion of replication initiation events in highly transcribed regions. Genome Res. 21: 1822-1832.

McBride D.J., Etemadmoghadam D., Cooke S.I., Alsop K., George J. et al. (2012). Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. J. Pathol. 227: 446-455.

Mesner L.D., Valsakumar V., Karnani N., Dutta A., Hamlin J. and Bekiranov S. (2010) Bubble-chip analysis of human origin distributions demonstrates on a genomic scale significant clustering into zones and significant association with transcription. Genome Res. 21: 377-389.

Mesner L.D., Valsakumar V., Cieslik M., Pickin R., Hamlin J.L. and Bekiranov S. (2013). Bubble-seq analysis of the human genome reveals distinct chromatin-mediated mechanisms for regulating early- and late-firing origins. Genome Res. (July 16 Epub ahead of print).

Neve R.M., Chin K., Fridlyand J., Yeh J., Baehner F.L., Fevr T., Clark L., Bayani N., Coppe J..P, Tong F., Speed T., Spellman P.T., DeVries S., Lapuk A., Wang N.J., Kuo W..L, Stilwell J.L., Pinkel D., Albertson D.G., Waldman F.M., McCormick F., Dickson R.B., Johnson M.D., Lippman M., Ethier S., Gazdar A. and Gray J.W. (2006). A

collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**: 515-527.

Nguyen V.Q., Co C. and Li J.J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. Nature 411: 1068-1073.

Nikolsky Y., Sviridov E., Yao J., Dosymbekov D., Ustyansky V., Kaznacheev V., Dezso Z., Mulvey L., Macconaill L.E., Winckler W., Serebryiskaya T., Nikolskaya T. and Polyak K. (2008). Genome-wide functional synergy between amplified and mutated genes in human breast cancer. Cancer Res. 68: 9532-9540.

Pan H., Deng, Y. and Pollard J.W. (2006). Progesterone blocks estrogen-induced DNA synthesis through the inhibition of replication licensing. Proc. Nat. Acad. Sci. 103: 14021-14026.

Payen C, Koszul R, Dujon B and Fischer G (2008). Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication based mechanisms. PLoS Genetio 4: e1000175.

Petropoulou C., Kotantaki P., Karamitros D. and Tara Viras S. (2008). Cdt1 and Geminin in cancer: markers or triggers of malignant transformation? Front. Biosci. 13: 4485-4494.

Quinlan A.R. and Hall I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26: 841-2.

Raveendranathan M., Chattopadhyay S., Bolon Y.T., Haworth J., Clarke D.J. and Bielinsky A.K. (2006). Genome-wide replication profiles of S-phase checkpoint mutants reveal fragile sites in yeast. EMBO J. 25: 3627-3639.

Ren B., Yu G., Tseng G.C., Cieply K., Gavel T. et al. (2006). MCM7 amplification and overexpression are associated with prostate cancer progression. Oncogene 25: 1090-1098.

Santarius T., Shipley J., Brewer D., Stratton M.R. and Cooper C.S. (2010). A census of amplified and overexpressed human cancer genes. Nature Rev. Cancer 10: 59-64.

Shi Y-K., Yu Y.P., Zhu Z-H., Han Y-C., Rec B., Nelson J.B. and Luo J-H. (2008). MCM7 interacts with androgen receptor. Amer. J. Path. 173: 1758-1767.

Shim J.W., Tan Q. and Gu L.-Q. (2009). Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity. Nucleic Acids Res. 37: 972–982.

Stark G.R. and Wahl G.M. (1984). Gene amplification. Annu Rev Biochem. 53: 447-491.

Teer J.K. and Dutta A. (2008). Human Cdt1 lacking the evolutionarily conserved region that interacts with MCM2-7 is capable of inducing re-replication. J. Biol. Chem. 283: 6817-6825.

Tao L., Dong Z., Leffak M., Zannis-Hadjopoulos M. and Price G. (2000). Major DNA replication initiation sites in the c-myc locus in human cells. J. Cell. Biochem. 78, 442-457.

Valenzuela M.S., Chen Y., Davis S., Yang F., Walker R.L., Bilke S., Lueders J., Martin M.M., Aladjem M.I., Massion P.P. and Meltzer P.S. (2011). Preferential localization of human origins of DNA replication at the 5'-ends of expressed genes and at evolutionarily conserved DNA sequences. PLoS One. 2011;6(5):e17308.

Welboren W.J., van Driel M.A., Janssen-Megens E.M., van Heeringen S.J., Sweep F.C., Span P.N. and Stunnenberg H.G. (2009). ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J. 28: 1418-1428.

Zhang Y., Liu T., Meyer C.A., Eeckhoute J., Johnson D.S., Bernstein B.E., Nusbaum C., Myers R.M., Brown M., Li W. and Liu X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol.9(9):R137.

Zhu W., Chen Y. and Dutta AQ. (2004). Rereplication by depletion of geminin is seen regardless of p53 status and activates a G2/M checkpoint. Mol. Cell. Biol. 24: 7140-7050.

Zhu W. and Dutta A. (2006). Activation of Fanconi anemia pathway in cells with re-replicated DNA. Cell Cycle 5: 2306-2309.

SUPPORTING DATA

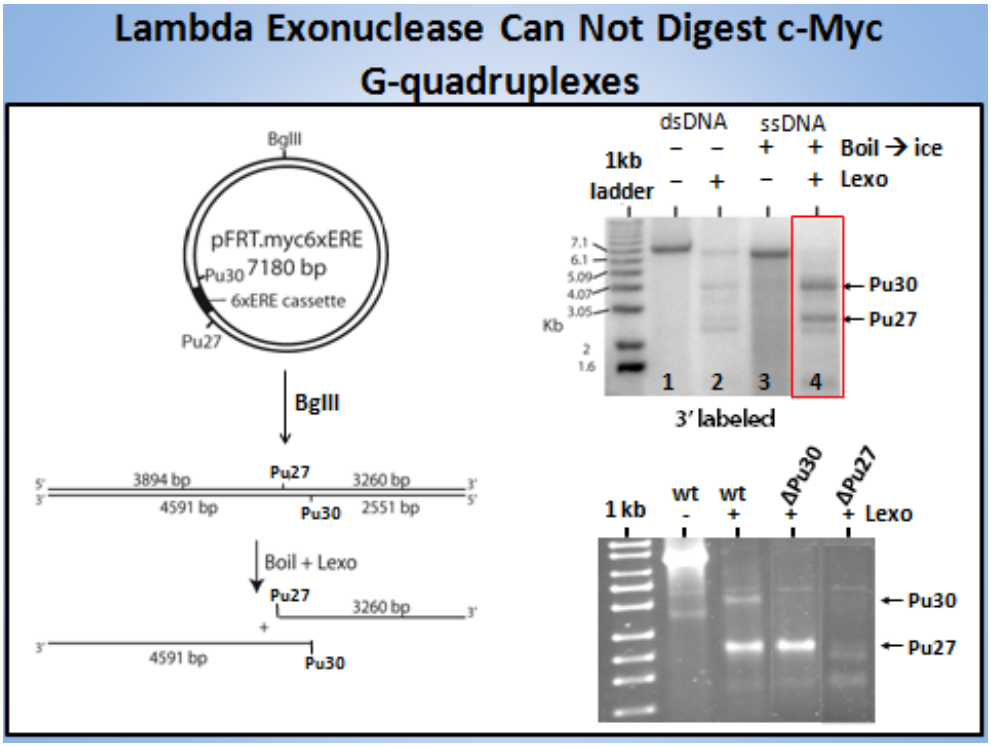


Figure 1. Lambda exonuclease (Lexo) cannot digest past G4 quadruplexes.

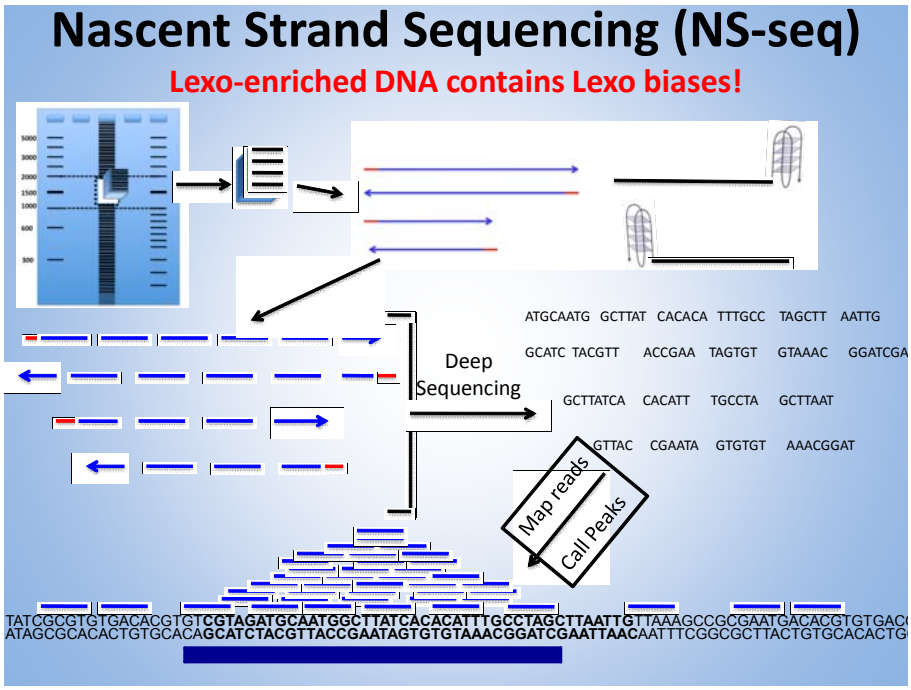


Figure 2. Lexo digested DNA is enriched in G-quadruplexes in addition to newly synthesized DNA.

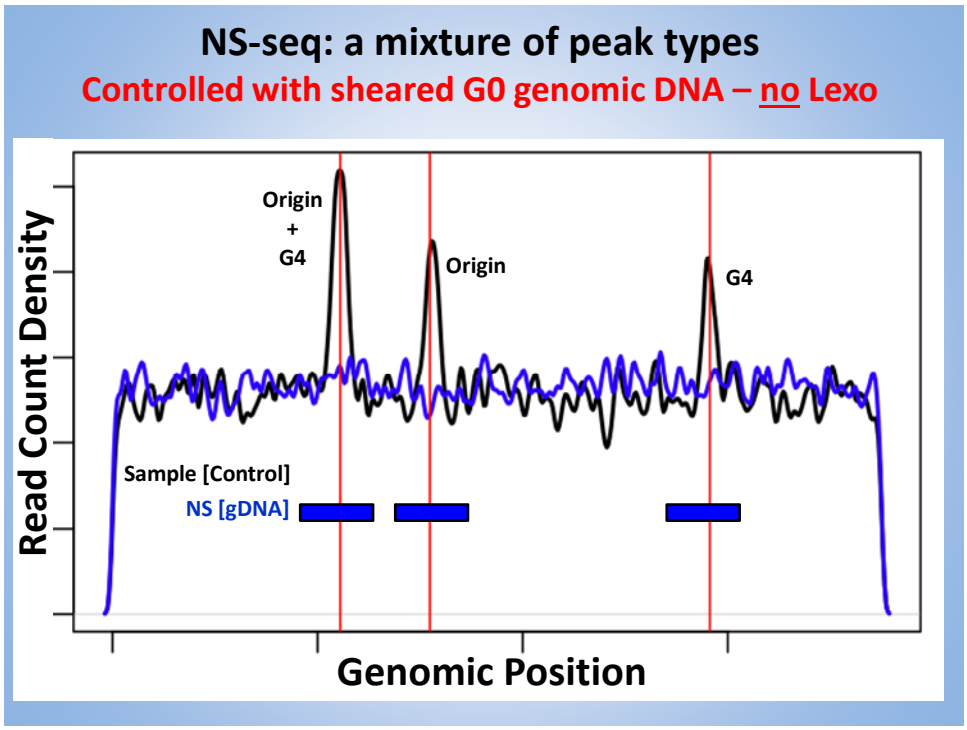


Figure 3. NS-seq will have G-quadruplex peaks in addition to replication origin peaks.

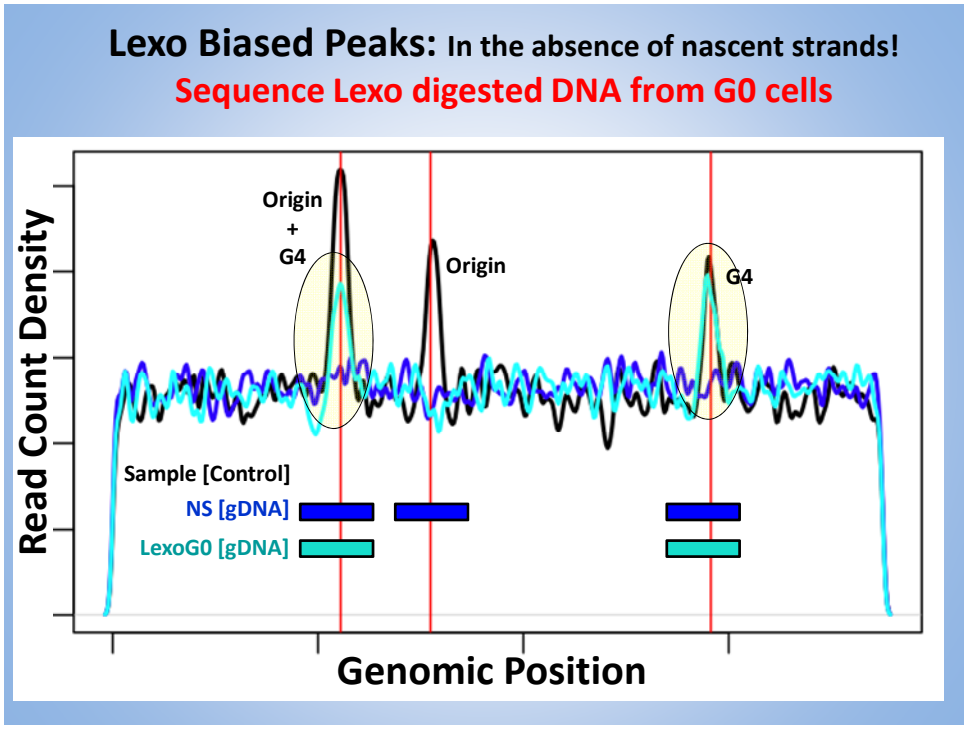


Figure 4. The genomic sequence profile of Lexo digested DNA from non-replicating (G0) cells will contain G4 peaks and replication origin peaks.

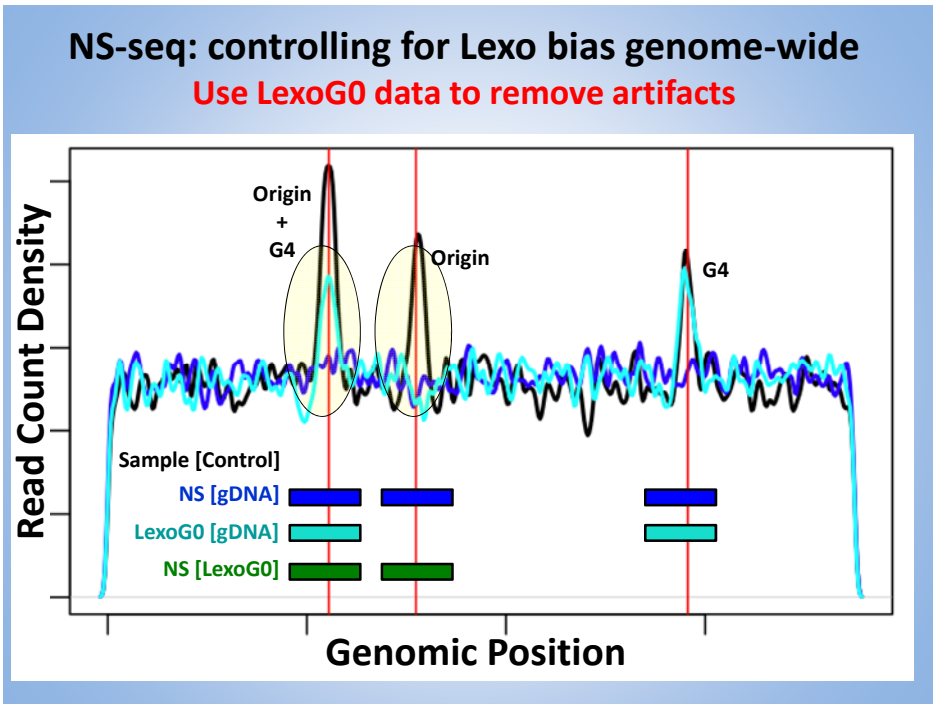


Figure 5. Lexo data from G0 cells can be used to correct Lexo data from replicating cells (NS-Seq).

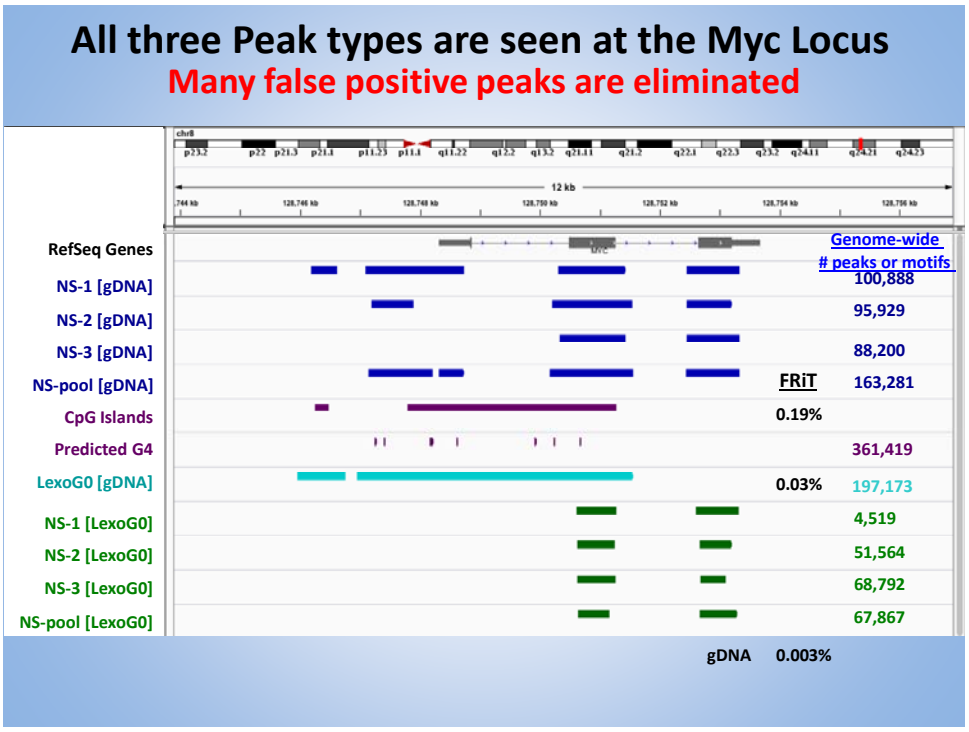


Figure 6. Correction by Lexo data from G0 cells reveals which of the three NS-Seq peaks at the Myc locus are replication origins.

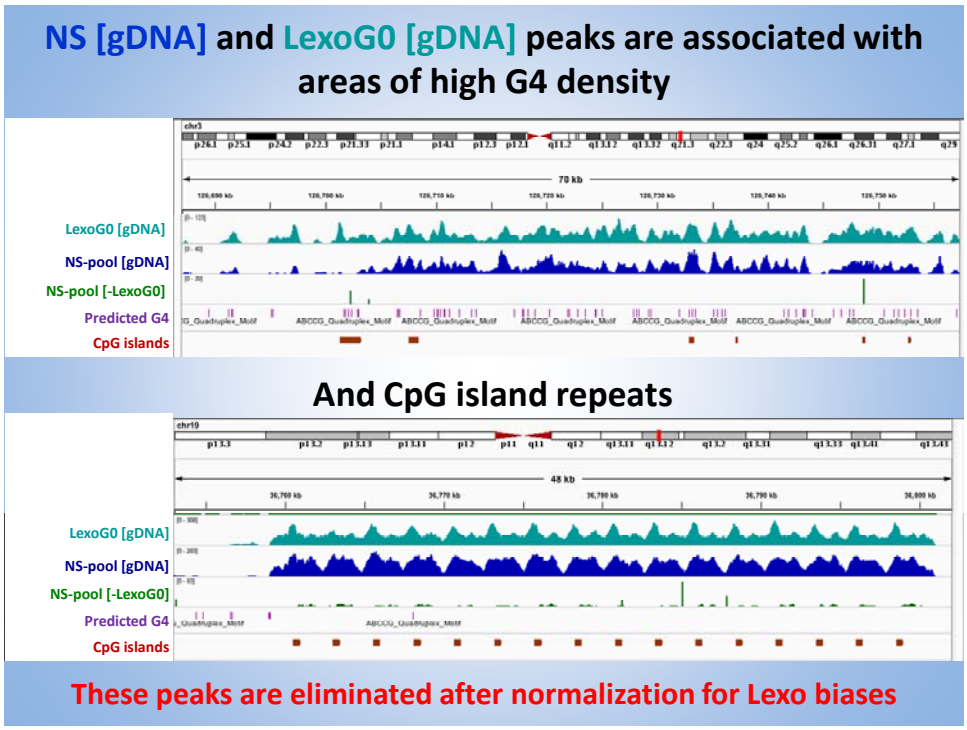


Figure 7. Peaks from genomic DNA (gDNA) and Lexo digested DNA from G0 cells correlate with G quadruplexes and CpG islands.

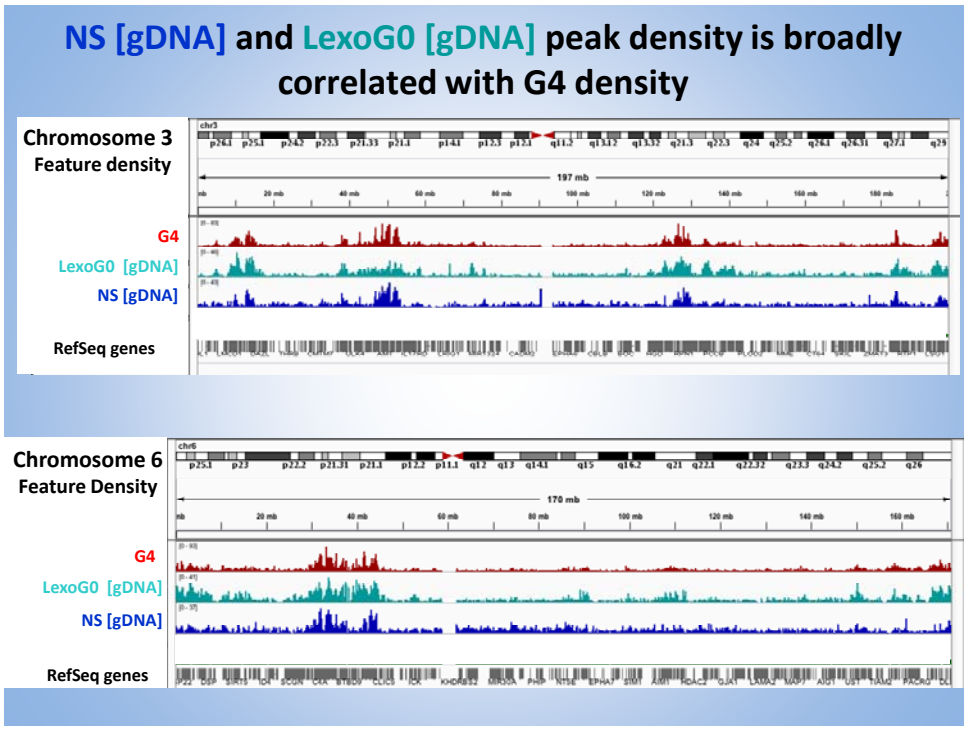


Figure 8. The correlation of gDNA and Lexo digested DNA from G0 cells correlate with G quadruplexes genome-wide (e.g., regions from chromosomes 3 and 6).

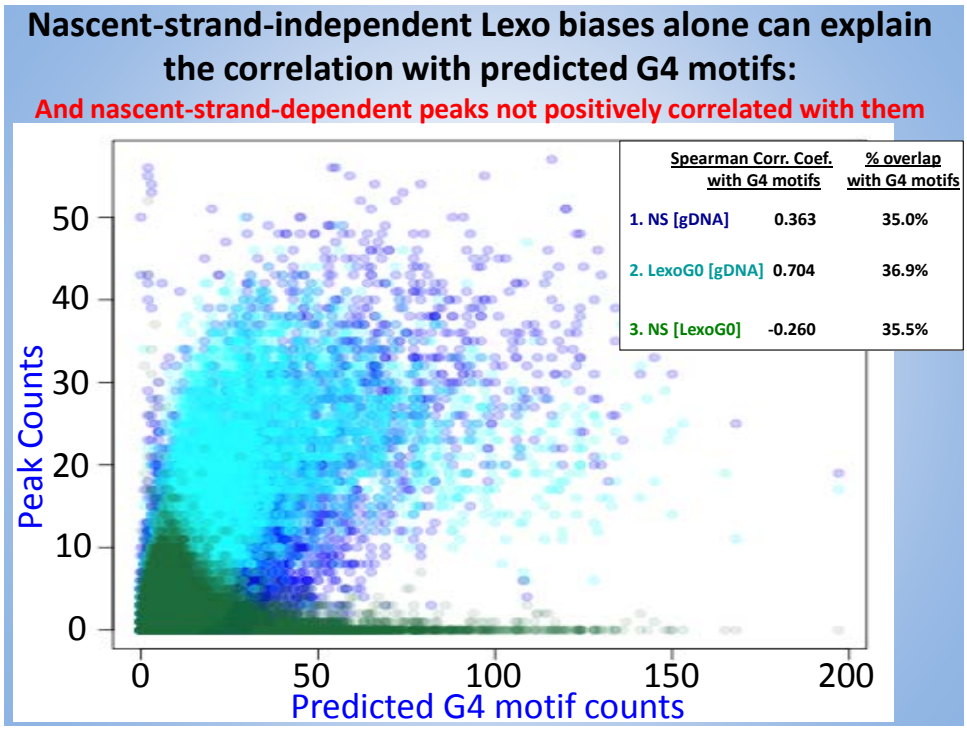


Figure 9. Nascent strand-independent Lexo biases can explain the correlation of replication origins with G-quadruplexes reported by others.

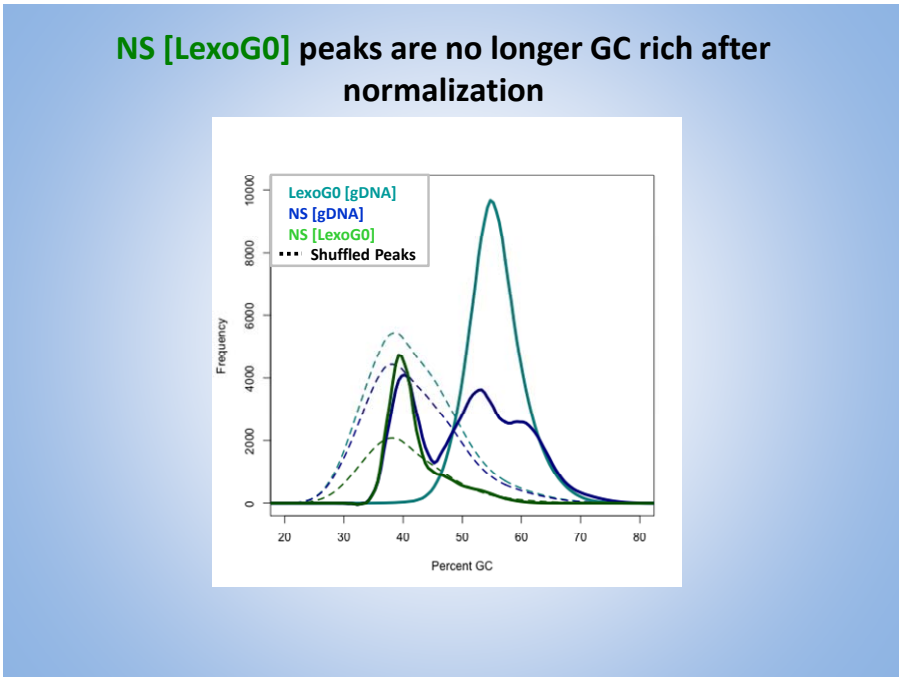


Figure 10. NS-seq peaks of replication origins after correction for Lexo biases are no longer GC rich.

Toward an Improved NS-seq Protocol

G-Quadruplex Stability is Cation Selective

Shim et al., (2009) *NAR* 37: 972-82:

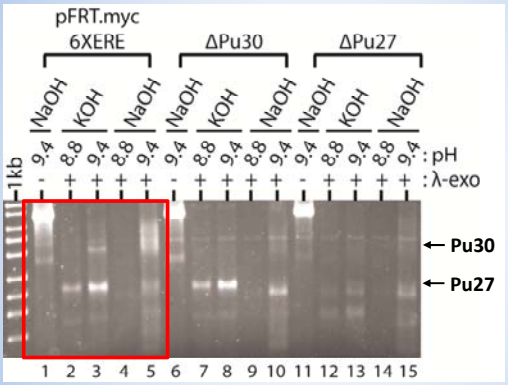
$K^+ > NH_4^+ > Ba^{2+} > Cs^+ > Na^+ > Li^+$

BUT…

- Traditional Lexo buffer is 67 mM glycine-KOH, pH 9.4
- We pioneered the practice of using 67 mM glycine-KOH pH 8.8 – protect RNA primers
- Test Lexo digestion of plasmid DNA with:
  67 mM glycine-NaOH pH 9.4
  67 mM glycine-NaOH pH 8.8

Figure 11. G-quadruplex stability depends on the cation (less stable in Na+ than K+).

c-Myc G-Quadruplexes are digested in Glycine-NaOH buffer



Concentration of KOH or NaOH @:

pH 8.8 ≈ 0.65 mM

pH 9.4 ≈ 1.95 mM

Simply switching to glycine-NaOH buffer may greatly reduce G4 Lexo bias!

Figure 12. Lexo digestion in Na+ buffer greatly reduces the Lexo G-quadruplex bias.

TABLE 1: Log of Breast Cancer Tumor Samples Received from R.I. Hospital

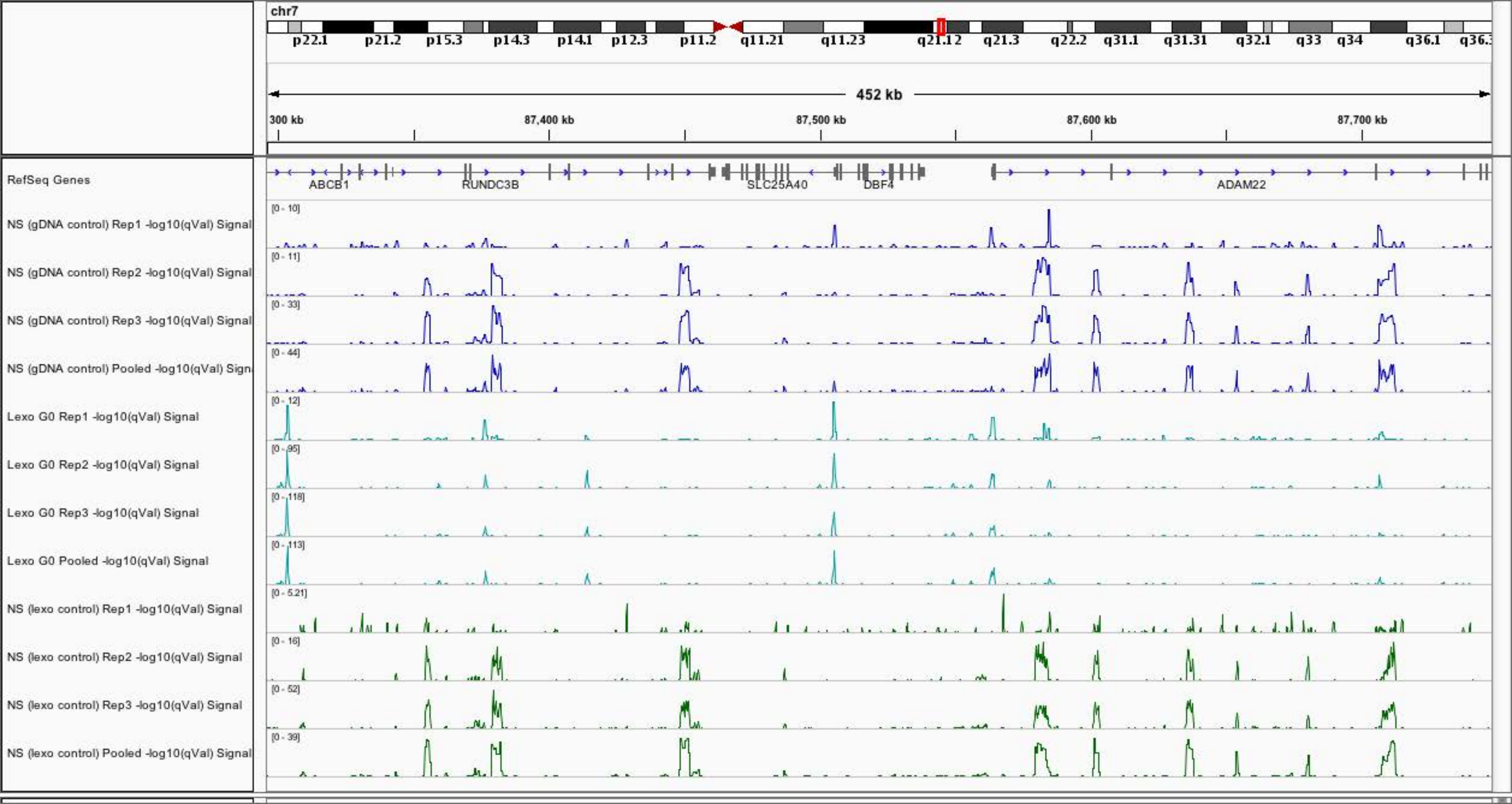
All tumor specimens were provided by Dr. Shamlal Mangray (Pathology Department, RI Hospital) and were frozen at -80 degrees. All samples were from female patients (identity unknown – coded by the Pathology Department) without neoadjuvent chemotherapy. Most were patients of Dr. Thersa Graves. The samples were 1.0-1.5 cm.

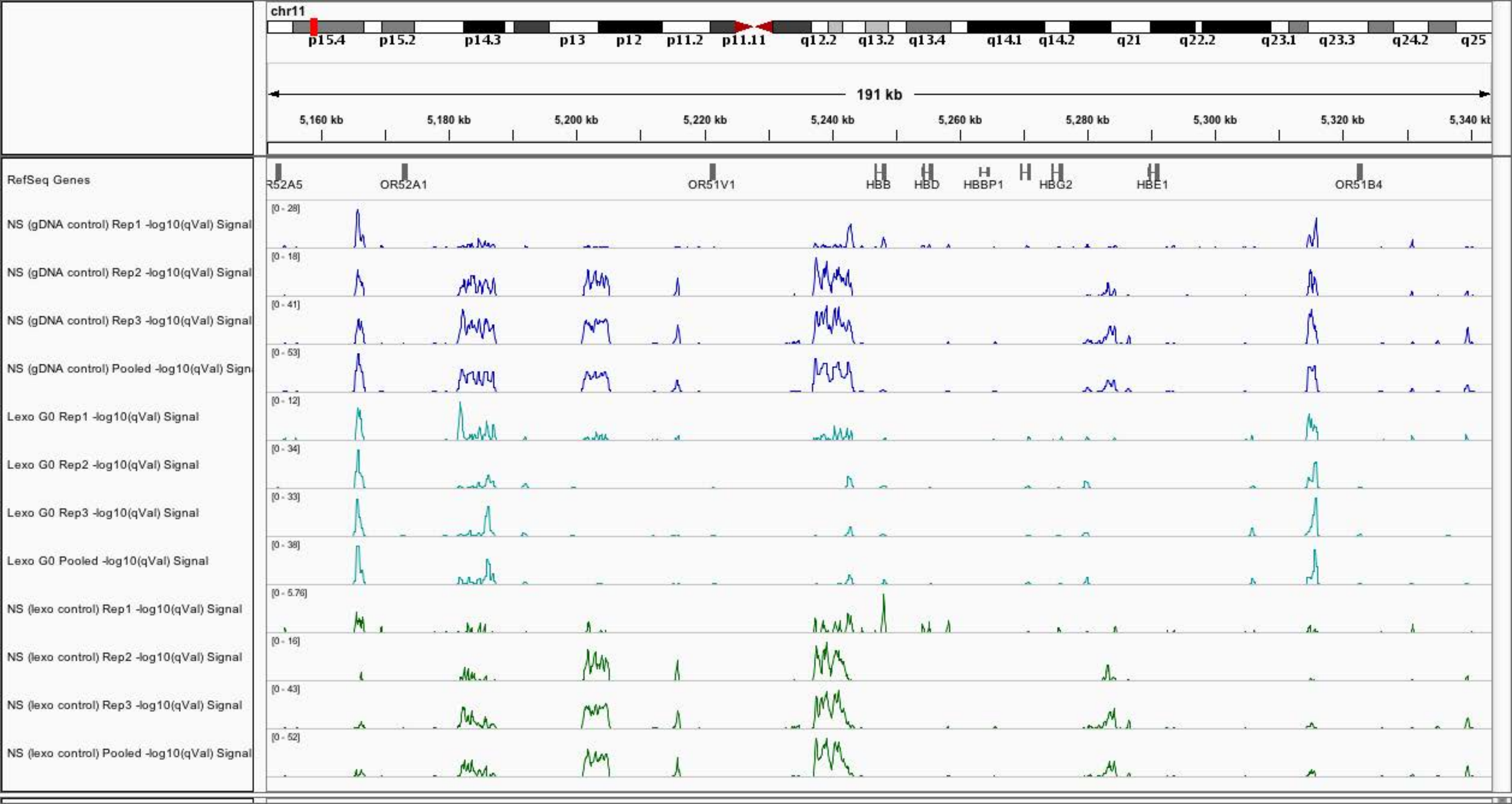
#	code	Date	ER	PR	HER2	Age	Comments
33	1	10/25/07	pos				used for H4 test1
34	2	11/13/07					also normal tissue
35	3	12/17/07	pos	pos	neg		also normal tissue
36	4	1/18/08			post-menopausal(PM)		normal tissue (tube J) Path # 213A/J
37	5						
38	6	1/18/08				45	used:C=H4 test (1 g) also normal tissue (tube F) Path # 375C/F
39	7						
40	8	1/28/08	pos	pos	neg	60(PM)	also normal tissue
41	9						
42	10 (SG5)	1/28/08	pos	pos	neg	49	used: H4/ER/mock(0.2 g) also normal tissue
43	11						
44	12 (SG6)	1/28/08	pos	pos	neg	41	also normal tissue
3	SG8	9/09	3+	2+	Neg	73	8 cm tumor, lymph node mets
5	SG10	9/09	3+	1+ (5-10%)	Neg	28	4 cm tumor, sentinel lymph node (SLN) micromets
6	SG11	9/09	3+	3+	Neg	54	2.5 cm tumor, no mets to SLN

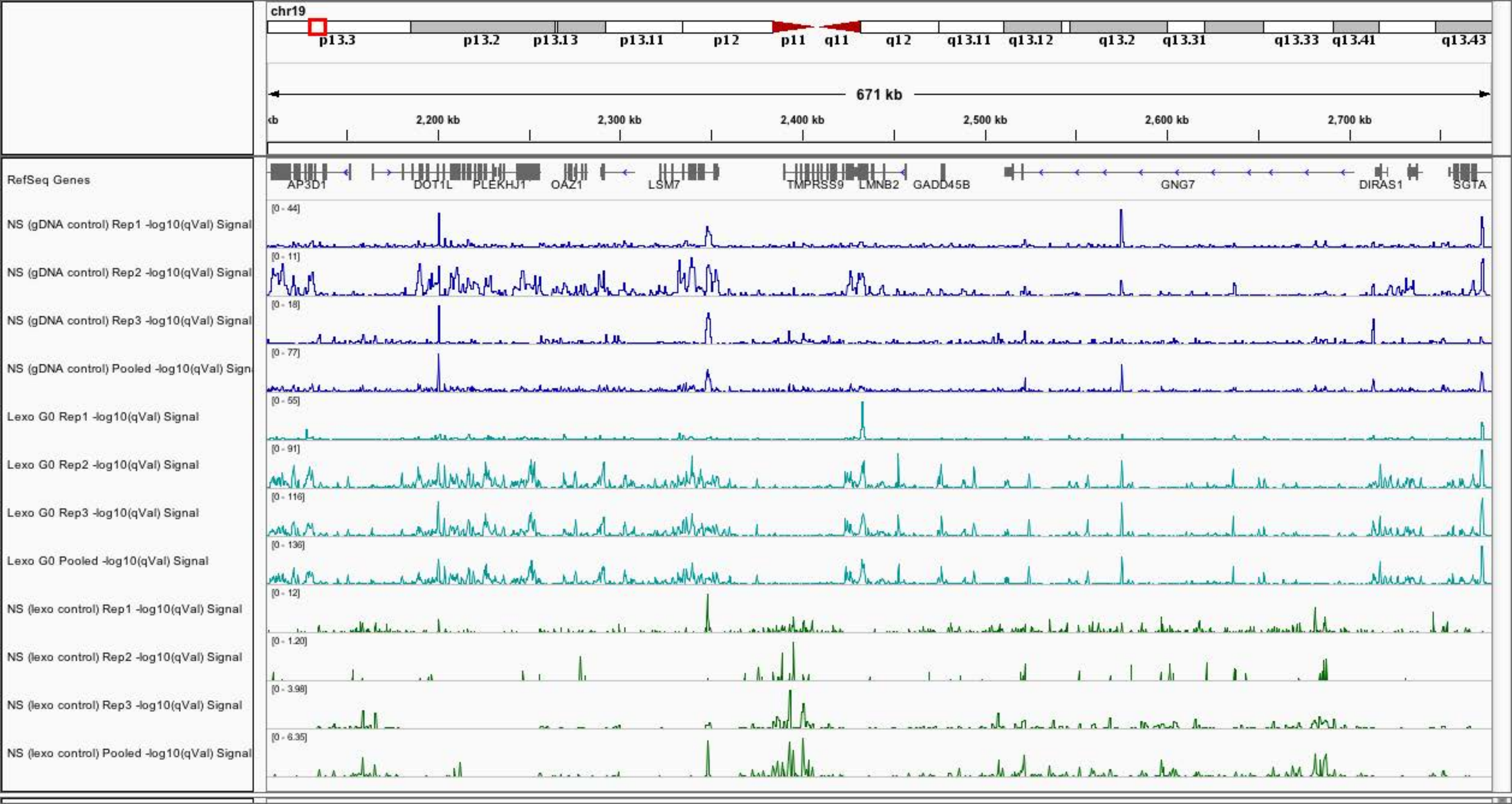
8	SG13	9/09	Neg	Neg	Neg	53	1.7 cm tumor, no mets to SLN
9	SG14	9/09	3+	3+	2+(FISH neg)	33	4 cm tumor no mets to SLN
10	SG15	9/09	3+	3+	2+(FISH neg)	66	2.4 cm tumor, No SLN sampling, FNA of node negative
11	SG16	9/09	Neg	Neg	2+(FISH neg)	80	1.1 cm tumor, no mets to SLN
12	SG17	9/09	3+	Neg	2+(FISH neg)	69	2.5 cm tumor, Axillary lymph node (ALN) negative
13	SG18	9/09	3+	3+	FISH neg	59	1.3 cm tumor, no mets to SLN
4	19	11/8/10	pos	pos	neg		also normal tissue
5	20	11/8/10	pos	pos	neg		also normal tissue
6	21	11/8/10	pos	pos	neg		also normal tissue
7	22	11/8/10	pos	pos	neg		also normal tissue
8	25	11/8/10	pos	pos	neg		also normal tissue
9	26	11/8/10	pos	pos	neg		also normal tissue
10	27	11/8/10	pos	pos	neg		also normal tissue
11	28	11/8/10	pos	pos	neg		also normal tissue
12	29	11/8/10	pos	pos	neg		also normal tissue
13	30	11/8/10	pos	pos	neg		also normal tissue
14	31	11/8/10	pos	pos	neg		also normal tissue

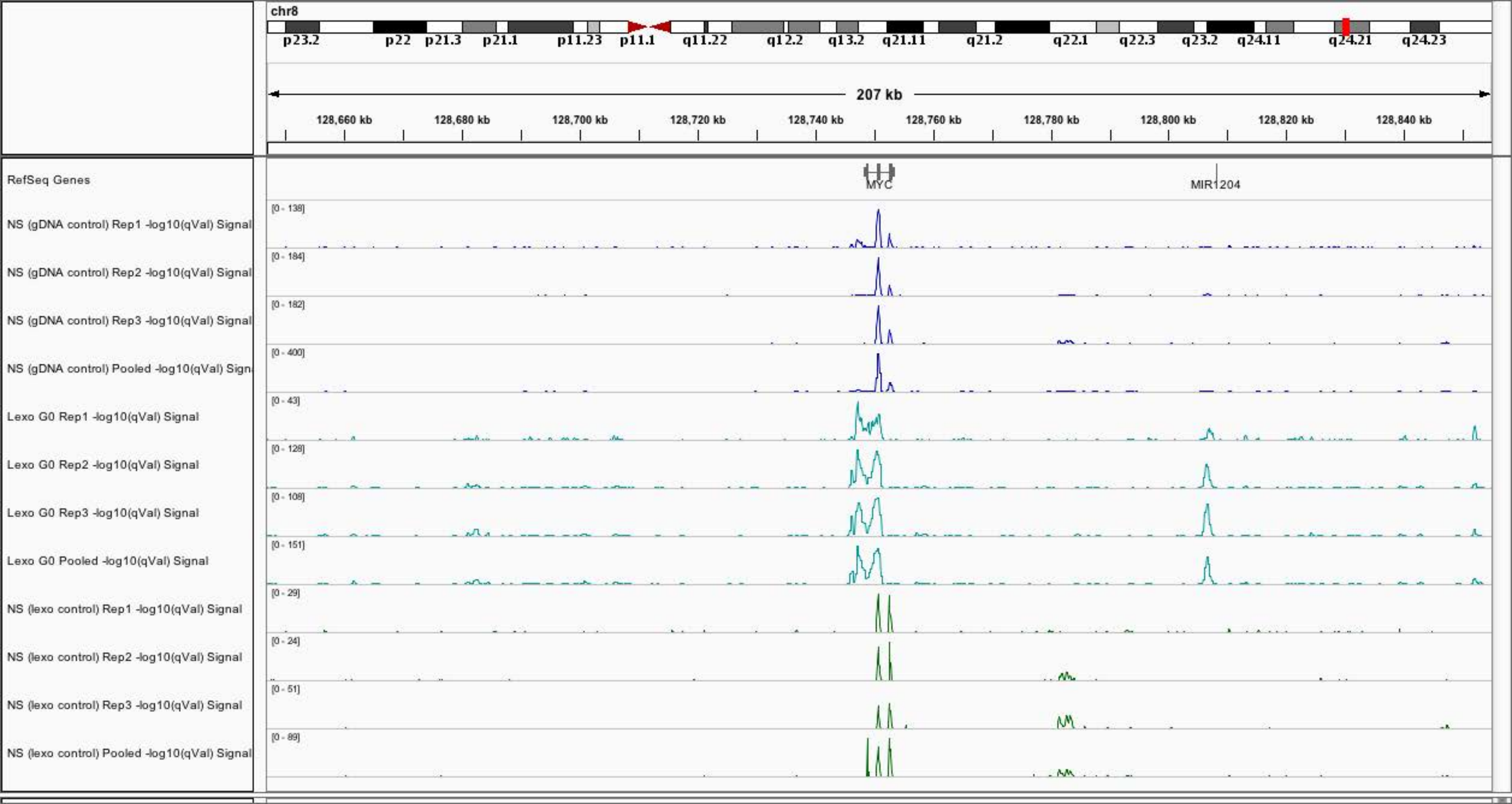
Appendix 1 Table of Contents:
(note – p. 1 of Appendix 1 is page 45 of the merged document, etc).

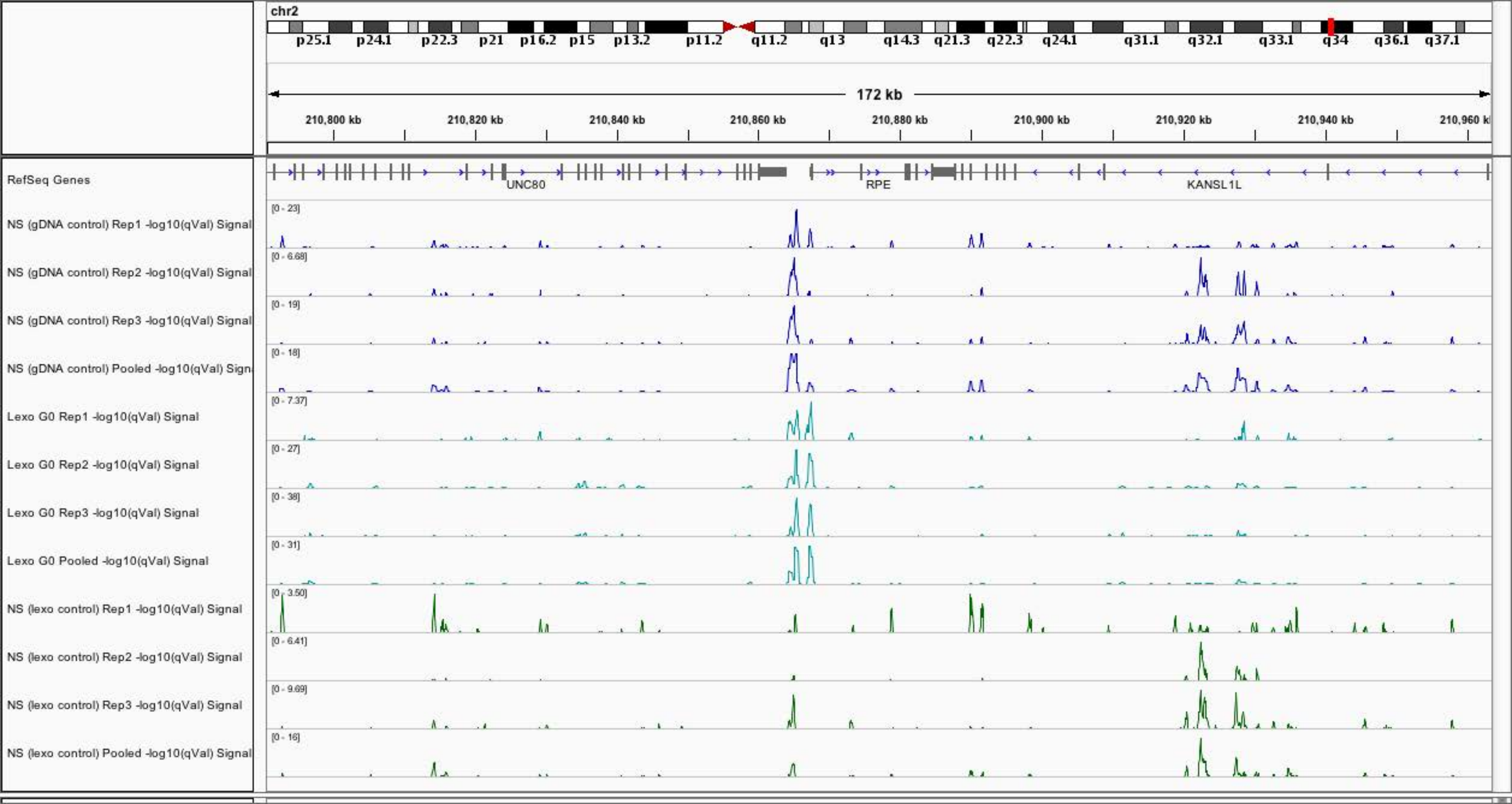
Appendix Table of Contents:

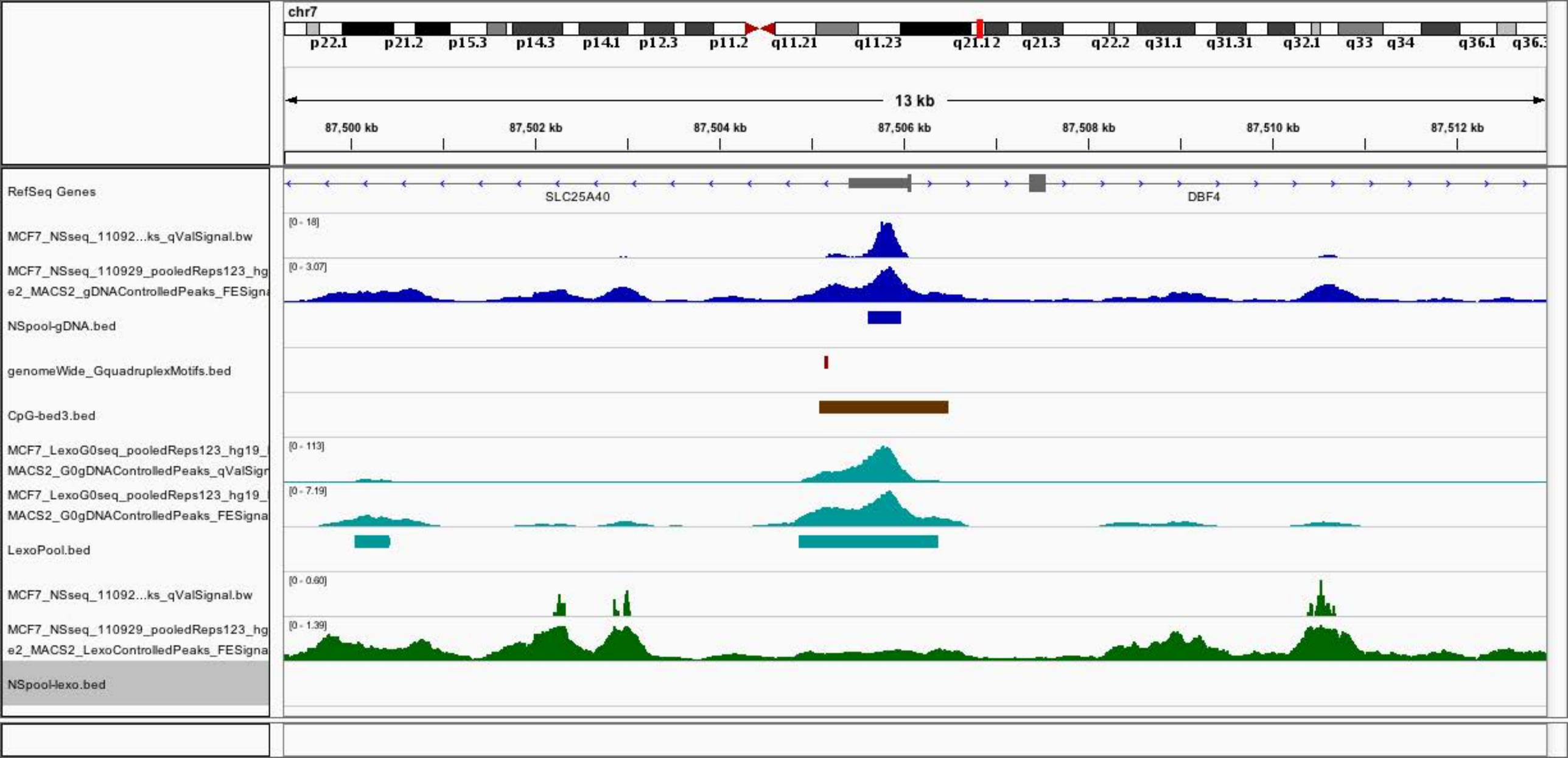

chr7
p22.1
p21.2
p15.3
p14.3
p14.1
p12.3
p11.2
q11.21
q11.23
q21.12
q21.3
q22.2
q31.1
q31.31
q32.1
q33
q34
q36.1
452 kb
300 kb
87,400 kb
87,500 kb
87,600 kb
87,700 kb
RefSeq Genes
ABCB1
RUNDC3B
SLC25A40
DBF4
ADAM22
NS (gDNA control) Rep1 -log10(qVal) Signal
[0 - 10]
NS (gDNA control) Rep2 -log10(qVal) Signal
[0 - 11]
NS (gDNA control) Rep3 -log10(qVal) Signal
[0 - 33]
NS (gDNA control) Pooled -log10(qVal) Sign
[0 - 44]
Lexo G0 Rep1 -log10(qVal) Signal
[0 - 12]
Lexo G0 Rep2 -log10(qVal) Signal
[0 - 95]
Lexo G0 Rep3 -log10(qVal) Signal
[0 - 118]
Lexo G0 Pooled -log10(qVal) Signal
[0 - 113]
NS (lexo control) Rep1 -log10(qVal) Signal
[0 - 5.21]
NS (lexo control) Rep2 -log10(qVal) Signal
[0 - 16]
NS (lexo control) Rep3 -log10(qVal) Signal
[0 - 52]
NS (lexo control) Pooled -log10(qVal) Signal
[0 - 39]

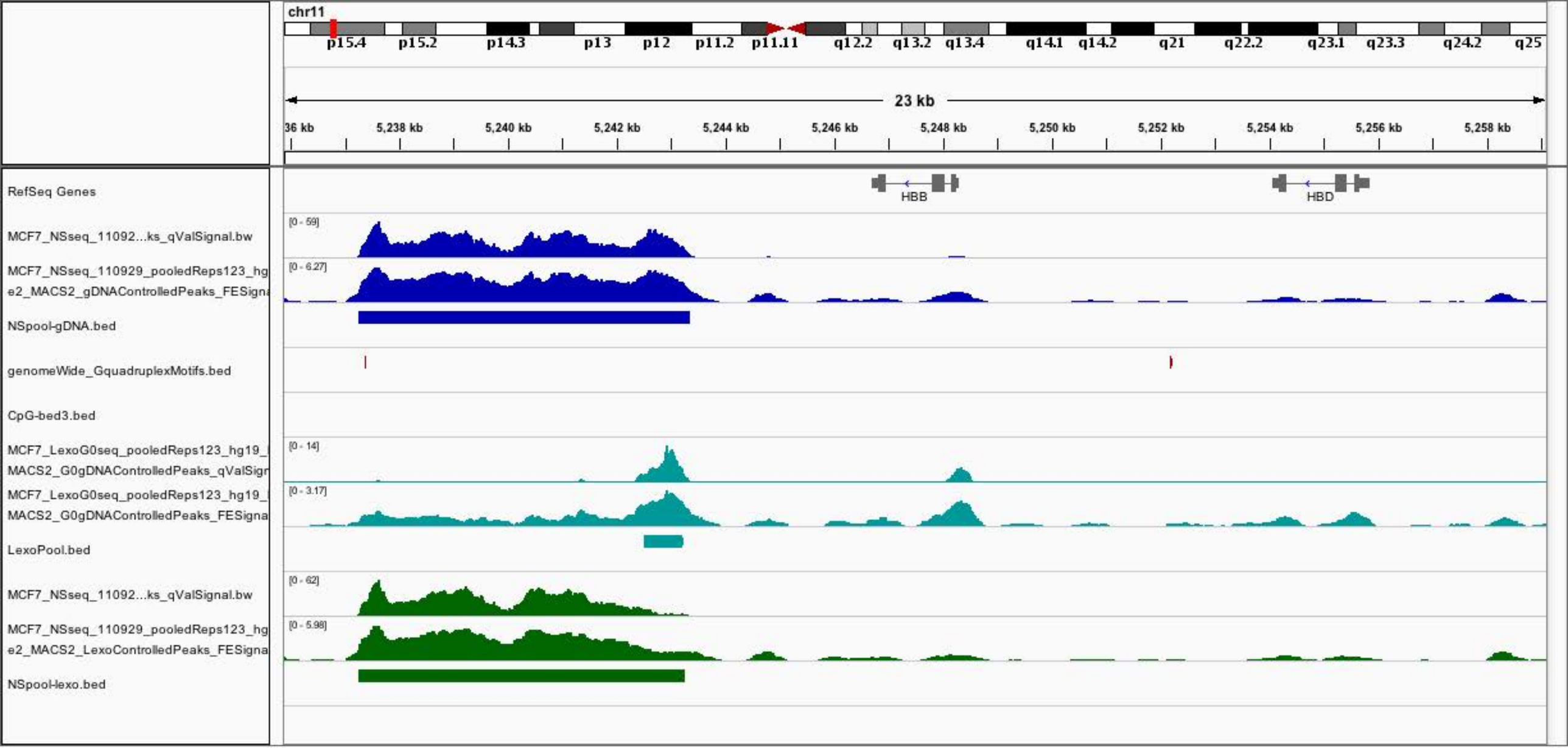

chr11
p15.4 p15.2 p14.3 p13 p12 p11.2 p11.11 q12.2 q13.2 q13.4 q14.1 q14.2 q21 q22.2 q23.1 q23.3 q24.2 q25
191 kb
5,160 kb 5,180 kb 5,200 kb 5,220 kb 5,240 kb 5,260 kb 5,280 kb 5,300 kb 5,320 kb 5,340 kb
RefSeq Genes
R52A5 OR52A1 OR51V1 HBB HBD HBBP1 HBG2 HBE1 OR51B4
NS (gDNA control) Rep1 -log10(qVal) Signal
[0 - 28]
NS (gDNA control) Rep2 -log10(qVal) Signal
[0 - 18]
NS (gDNA control) Rep3 -log10(qVal) Signal
[0 - 41]
NS (gDNA control) Pooled -log10(qVal) Sign
[0 - 53]
Lexo G0 Rep1 -log10(qVal) Signal
[0 - 12]
Lexo G0 Rep2 -log10(qVal) Signal
[0 - 34]
Lexo G0 Rep3 -log10(qVal) Signal
[0 - 33]
Lexo G0 Pooled -log10(qVal) Signal
[0 - 38]
NS (lexo control) Rep1 -log10(qVal) Signal
[0 - 5.76]
NS (lexo control) Rep2 -log10(qVal) Signal
[0 - 16]
NS (lexo control) Rep3 -log10(qVal) Signal
[0 - 43]
NS (lexo control) Pooled -log10(qVal) Signal
[0 - 52]

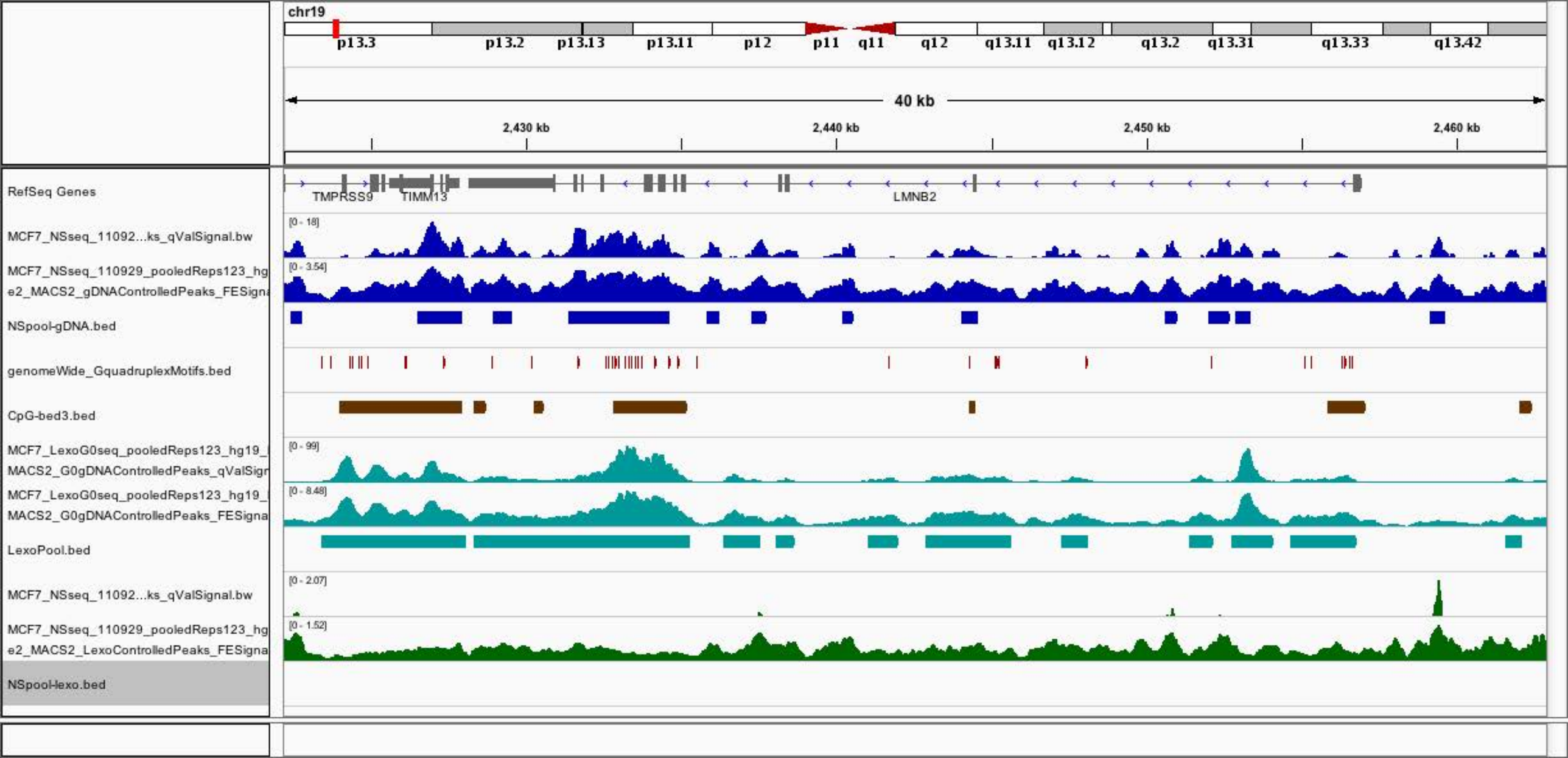

chr19
p13.3
p13.2
p13.13
p13.11
p12
p11
q11
q12
q13.11
q13.12
q13.2
q13.31
q13.33
q13.41
q13.43
671 kb
2,200 kb
2,300 kb
2,400 kb
2,500 kb
2,600 kb
2,700 kb
RefSeq Genes
AP3D1
DOT1L
PLEKHJ1
OAZ1
LSM7
TMPRSS9
LMNB2
GADD45B
GNG7
DIRAS1
SGTA
NS (gDNA control) Rep1 -log10(qVal) Signal
[0 - 44]
NS (gDNA control) Rep2 -log10(qVal) Signal
[0 - 11]
NS (gDNA control) Rep3 -log10(qVal) Signal
[0 - 18]
NS (gDNA control) Pooled -log10(qVal) Sign
[0 - 77]
Lexo G0 Rep1 -log10(qVal) Signal
[0 - 55]
Lexo G0 Rep2 -log10(qVal) Signal
[0 - 91]
Lexo G0 Rep3 -log10(qVal) Signal
[0 - 116]
Lexo G0 Pooled -log10(qVal) Signal
[0 - 136]
NS (lexo control) Rep1 -log10(qVal) Signal
[0 - 12]
NS (lexo control) Rep2 -log10(qVal) Signal
[0 - 1.20]
NS (lexo control) Rep3 -log10(qVal) Signal
[0 - 3.98]
NS (lexo control) Pooled -log10(qVal) Signal
[0 - 6.35]

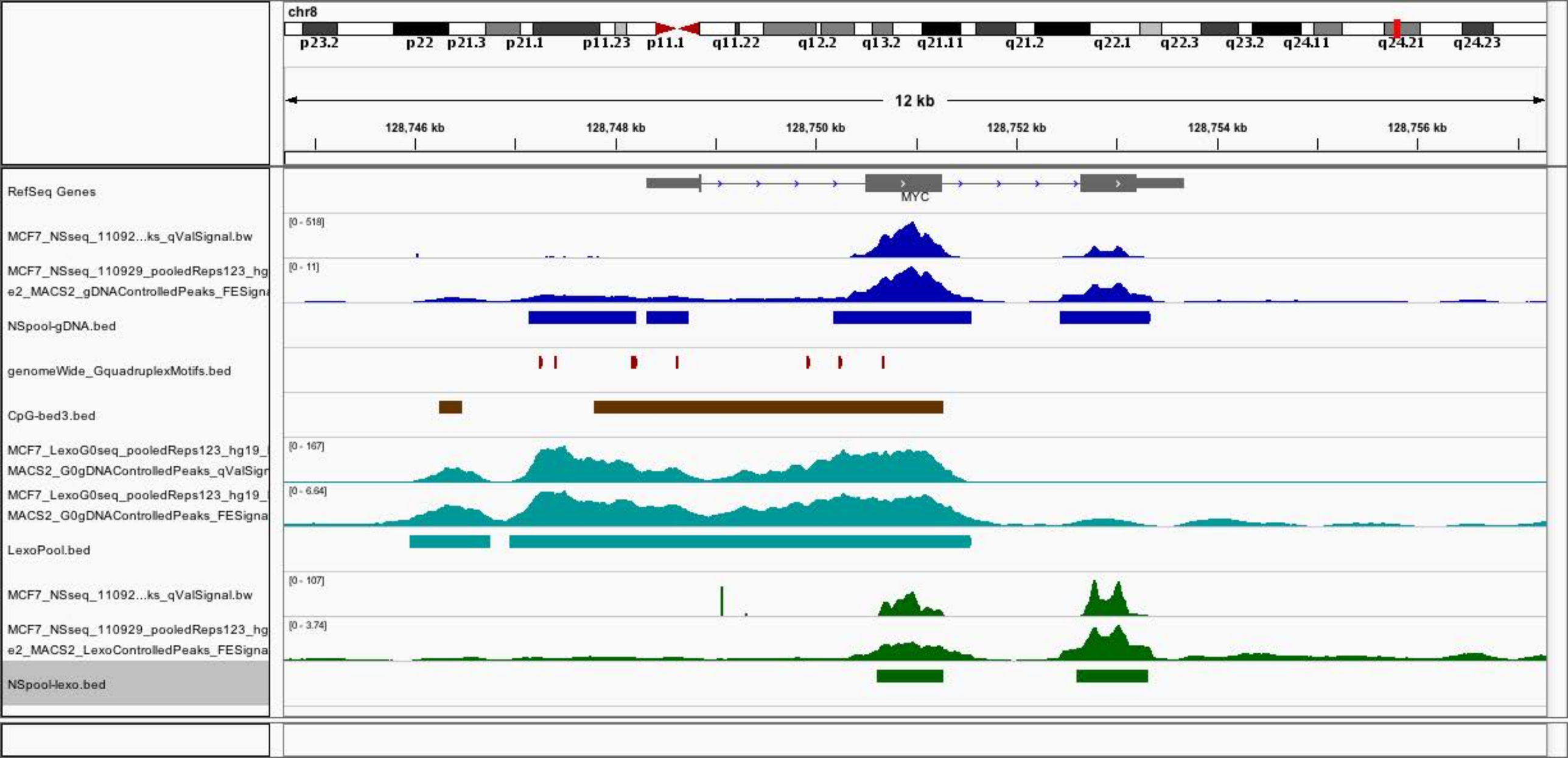

chr8
p23.2
p22
p21.3
p21.1
p11.23
p11.1
q11.22
q12.2
q13.2
q21.11
q21.2
q22.1
q22.3
q23.2
q24.11
q24.21
q24.23
207 kb
128,660 kb
128,680 kb
128,700 kb
128,720 kb
128,740 kb
128,760 kb
128,780 kb
128,800 kb
128,820 kb
128,840 kb
RefSeq Genes
MYC
MIR1204
NS (gDNA control) Rep1 -log10(qVal) Signal
[0 - 138]
NS (gDNA control) Rep2 -log10(qVal) Signal
[0 - 184]
NS (gDNA control) Rep3 -log10(qVal) Signal
[0 - 182]
NS (gDNA control) Pooled -log10(qVal) Sign
[0 - 400]
Lexo G0 Rep1 -log10(qVal) Signal
[0 - 43]
Lexo G0 Rep2 -log10(qVal) Signal
[0 - 128]
Lexo G0 Rep3 -log10(qVal) Signal
[0 - 108]
Lexo G0 Pooled -log10(qVal) Signal
[0 - 151]
NS (lexo control) Rep1 -log10(qVal) Signal
[0 - 29]
NS (lexo control) Rep2 -log10(qVal) Signal
[0 - 24]
NS (lexo control) Rep3 -log10(qVal) Signal
[0 - 51]
NS (lexo control) Pooled -log10(qVal) Signal
[0 - 89]

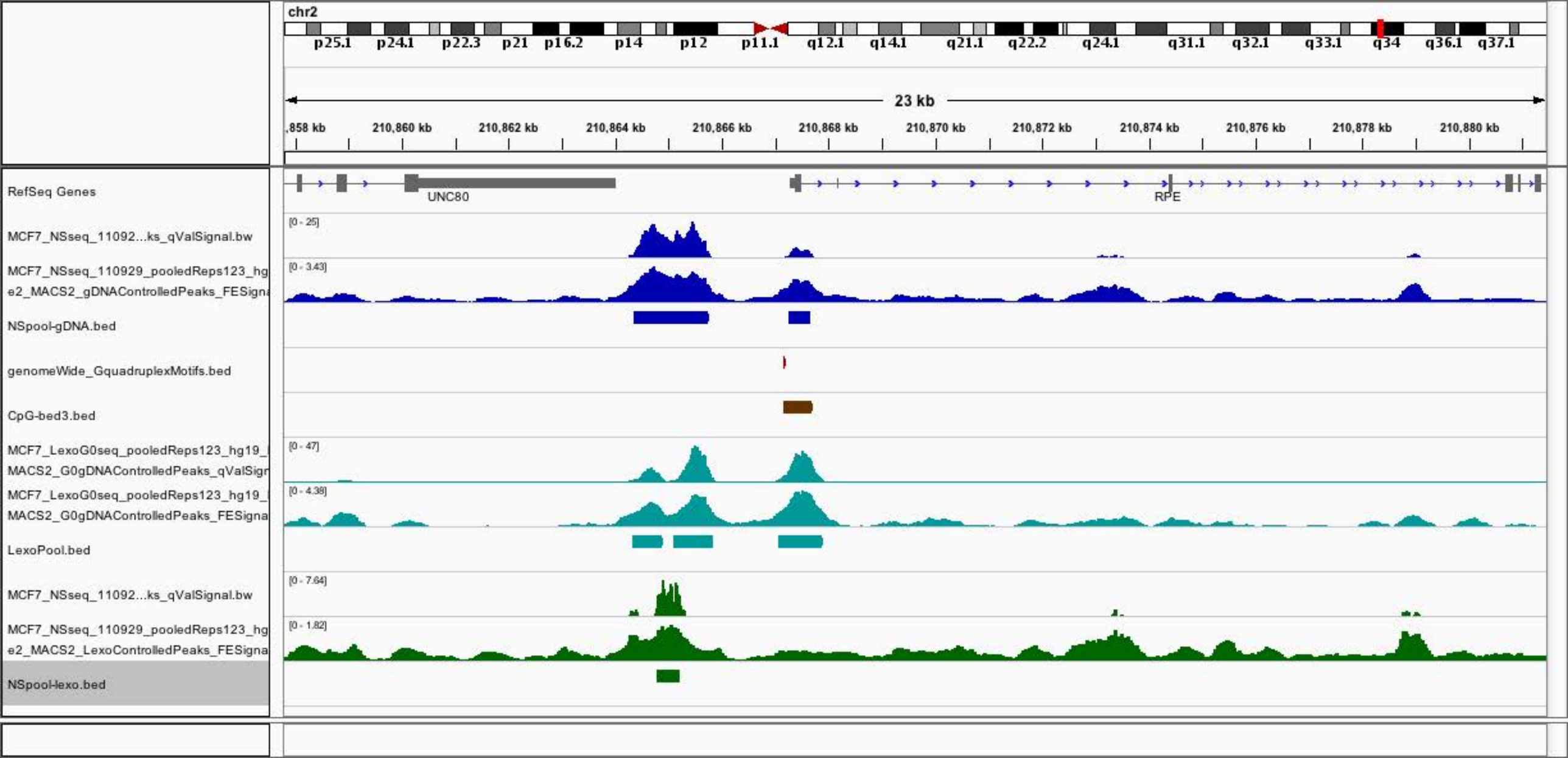

chr2
p25.1 p24.1 p22.3 p21 p16.2 p15 p13.2 p11.2 q11.2 q13 q14.3 q21.3 q22.3 q24.1 q31.1 q32.1 q33.1 q34 q36.1 q37.1
172 kb
210,800 kb
210,820 kb
210,840 kb
210,860 kb
210,880 kb
210,900 kb
210,920 kb
210,940 kb
RefSeq Genes
UNC80
RPE
KANSL1L
NS (gDNA control) Rep1 -log10(qVal) Signal
[0 - 23]
NS (gDNA control) Rep2 -log10(qVal) Signal
[0 - 6.68]
NS (gDNA control) Rep3 -log10(qVal) Signal
[0 - 19]
NS (gDNA control) Pooled -log10(qVal) Sign
[0 - 18]
Lexo G0 Rep1 -log10(qVal) Signal
[0 - 7.37]
Lexo G0 Rep2 -log10(qVal) Signal
[0 - 27]
Lexo G0 Rep3 -log10(qVal) Signal
[0 - 38]
Lexo G0 Pooled -log10(qVal) Signal
[0 - 31]
NS (lexo control) Rep1 -log10(qVal) Signal
[0 - 3.50]
NS (lexo control) Rep2 -log10(qVal) Signal
[0 - 6.41]
NS (lexo control) Rep3 -log10(qVal) Signal
[0 - 9.69]
NS (lexo control) Pooled -log10(qVal) Signal
[0 - 16]

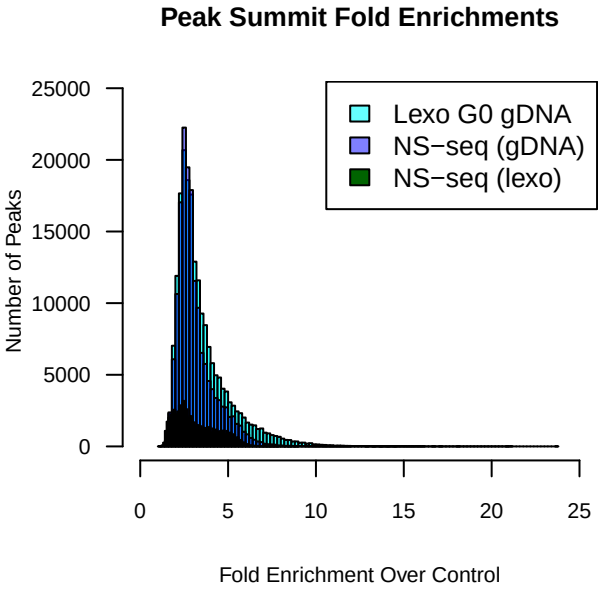

chr7
p22.1 p21.2 p15.3 p14.3 p14.1 p12.3 p11.2 q11.21 q11.23 q21.12 q21.3 q22.2 q31.1 q31.31 q32.1 q33 q34 q36.1 q36.3
13 kb
87,500 kb
87,502 kb
87,504 kb
87,506 kb
87,508 kb
87,510 kb
87,512 kb
RefSeq Genes
SLC25A40
DBF4
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 18]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_gDNAControlledPeaks_FESigna
[0 - 3.07]
NSpool-gDNA.bed
genomeWide_GquadruplexMotifs.bed
CpG-bed3.bed
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_qValSign
[0 - 113]
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_FESigna
[0 - 7.19]
LexoPool.bed
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 0.60]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_LexoControlledPeaks_FESigna
[0 - 1.39]
NSpool-lexo.bed

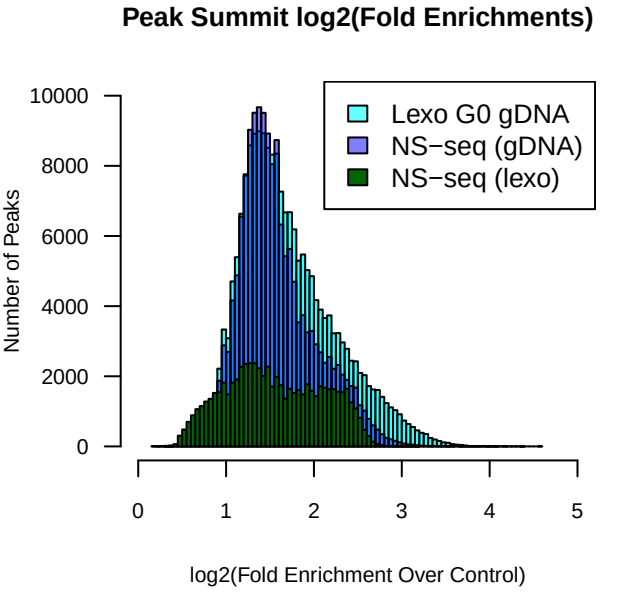

chr11
p15.4
p15.2
p14.3
p13
p12
p11.2
p11.11
q12.2
q13.2
q13.4
q14.1
q14.2
q21
q22.2
q23.1
q23.3
q24.2
q25
23 kb
36 kb
5,238 kb
5,240 kb
5,242 kb
5,244 kb
5,246 kb
5,248 kb
5,250 kb
5,252 kb
5,254 kb
5,256 kb
5,258 kb
RefSeq Genes
HBB
HBD
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 59]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_gDNAControlledPeaks_FESigna
[0 - 6.27]
NSpool-gDNA.bed
genomeWide_GquadruplexMotifs.bed
CpG-bed3.bed
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_qValSign
[0 - 14]
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_FESigna
[0 - 3.17]
LexoPool.bed
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 62]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_LexoControlledPeaks_FESigna
[0 - 5.98]
NSpool-lexo.bed

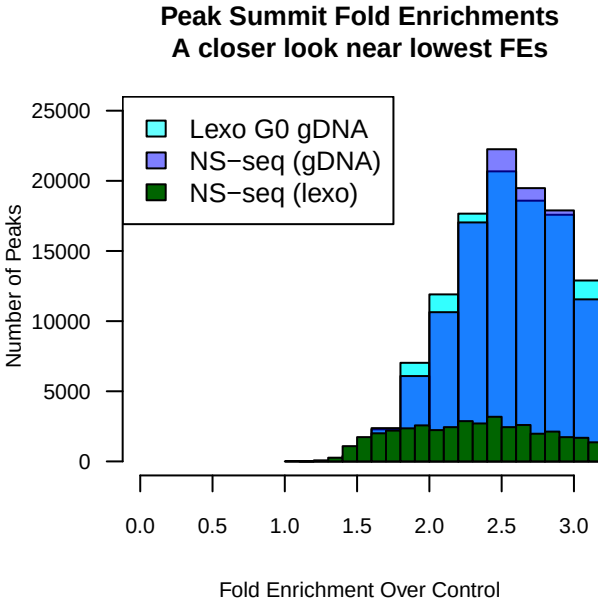

chr19
p13.3
p13.2
p13.13
p13.11
p12
p11
q11
q12
q13.11
q13.12
q13.2
q13.31
q13.33
q13.42
40 kb
2,430 kb
2,440 kb
2,450 kb
2,460 kb
RefSeq Genes
TMPRSS9
TIMM13
LMNB2
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 18]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_gDNAControlledPeaks_FESigna
[0 - 3.54]
NSpool-gDNA.bed
genomeWide_GquadruplexMotifs.bed
CpG-bed3.bed
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_qValSign
[0 - 99]
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_FESigna
[0 - 8.48]
LexoPool.bed
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 2.07]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_LexoControlledPeaks_FESigna
[0 - 1.52]
NSpool-lexo.bed

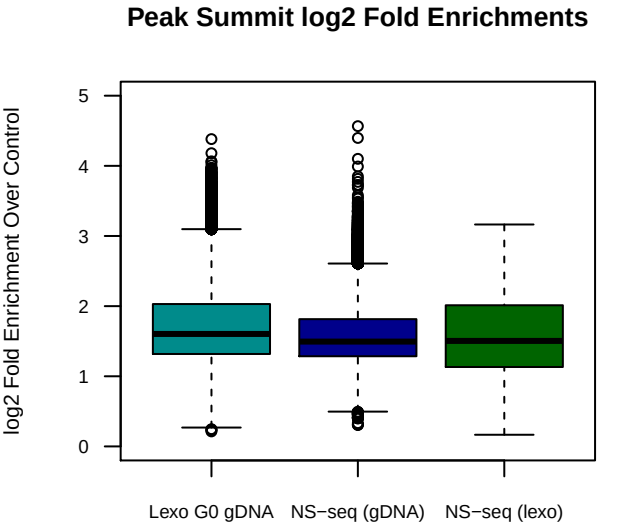

chr8
p23.2
p22
p21.3
p21.1
p11.23
p11.1
q11.22
q12.2
q13.2
q21.11
q21.2
q22.1
q22.3
q23.2
q24.11
q24.21
q24.23
12 kb
128,746 kb
128,748 kb
128,750 kb
128,752 kb
128,754 kb
128,756 kb
RefSeq Genes
MYC
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 518]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_gDNAControlledPeaks_FESigna
[0 - 11]
NSpool-gDNA.bed
genomeWide_GquadruplexMotifs.bed
CpG-bed3.bed
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_qValSign
[0 - 167]
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_FESigna
[0 - 6.64]
LexoPool.bed
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 107]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_LexoControlledPeaks_FESigna
[0 - 3.74]
NSpool-lexo.bed

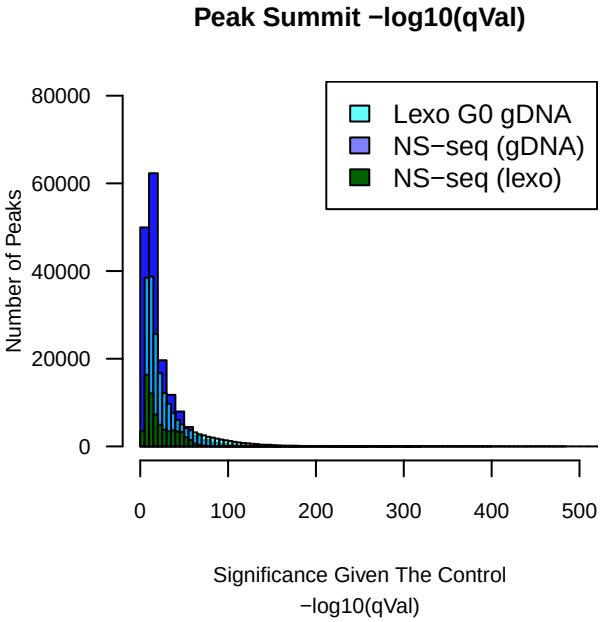

chr2
p25.1 p24.1 p22.3 p21 p16.2 p14 p12 p11.1 q12.1 q14.1 q21.1 q22.2 q24.1 q31.1 q32.1 q33.1 q34 q36.1 q37.1
23 kb
,858 kb
210,860 kb
210,862 kb
210,864 kb
210,866 kb
210,868 kb
210,870 kb
210,872 kb
210,874 kb
210,876 kb
210,878 kb
210,880 kb
RefSeq Genes
UNC80
RPE
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 25]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_gDNAControlledPeaks_FESigna
[0 - 3.43]
NSpool-gDNA.bed
genomeWide_GquadruplexMotifs.bed
CpG-bed3.bed
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_qValSign
[0 - 47]
MCF7_LexoG0seq_pooledReps123_hg19_
MACS2_G0gDNAControlledPeaks_FESigna
[0 - 4.38]
LexoPool.bed
MCF7_NSseq_11092...ks_qValSignal.bw
[0 - 7.64]
MCF7_NSseq_110929_pooledReps123_hg
e2_MACS2_LexoControlledPeaks_FESigna
[0 - 1.82]
NSpool-lexo.bed

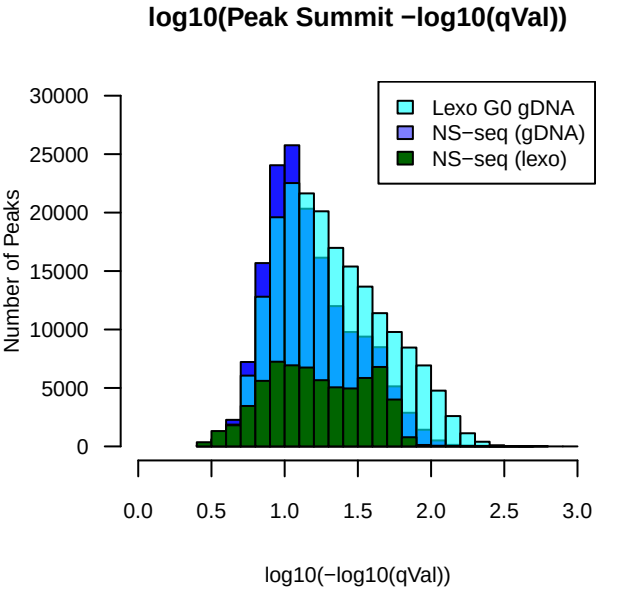

Peak Summit Fold Enrichments
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
25000
20000
15000
10000
5000
0
0
5
10
15
20
25
Fold Enrichment Over Control

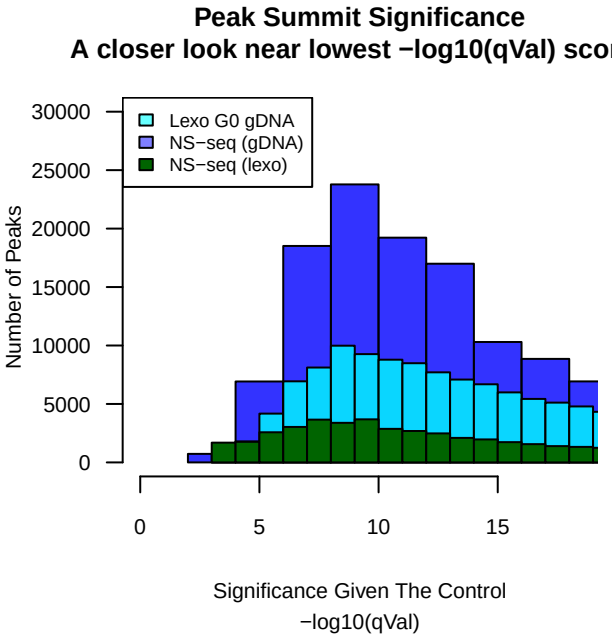

Peak Summit log2(Fold Enrichments)
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
10000
8000
6000
4000
2000
0
0
1
2
3
4
5
log2(Fold Enrichment Over Control)

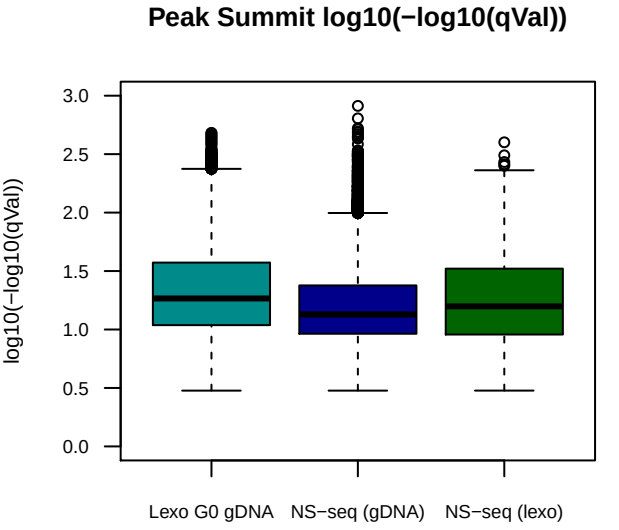

Peak Summit Fold Enrichments
A closer look near lowest FEs
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
25000
20000
15000
10000
5000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Fold Enrichment Over Control

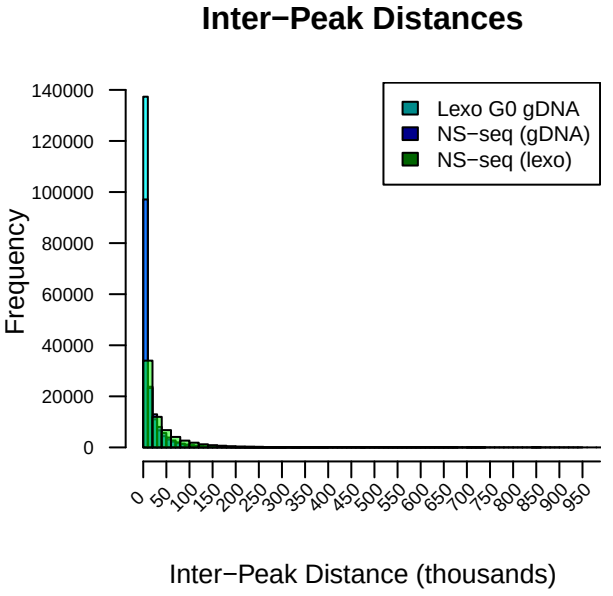

Peak Summit log2 Fold Enrichments
log2 Fold Enrichment Over Control
5
4
3
2
1
0
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)

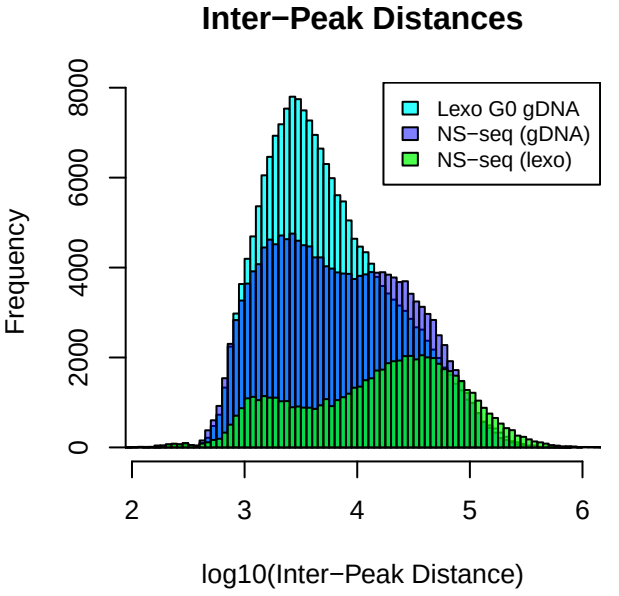

Peak Summit −log10(qVal)
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
80000
60000
40000
20000
0
0
100
200
300
400
500
Significance Given The Control
−log10(qVal)

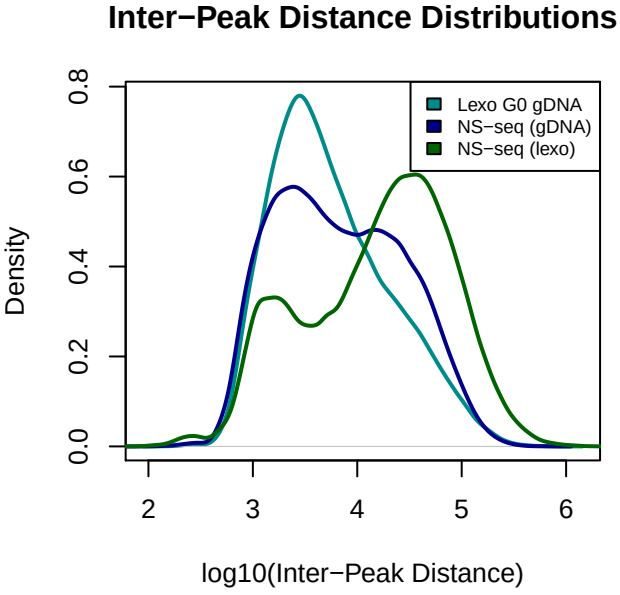

log10(Peak Summit −log10(qVal))
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
30000
25000
20000
15000
10000
5000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
log10(−log10(qVal))

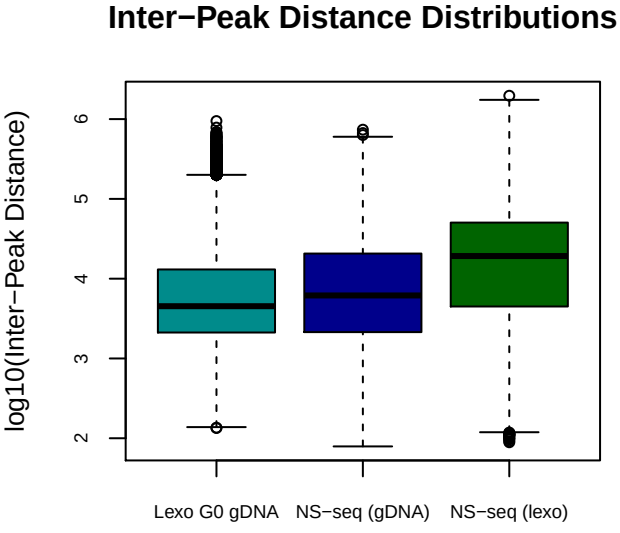

Peak Summit Significance
A closer look near lowest −log10(qVal) scores
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
30000
25000
20000
15000
10000
5000
0
0
5
10
15
Significance Given The Control
−log10(qVal)

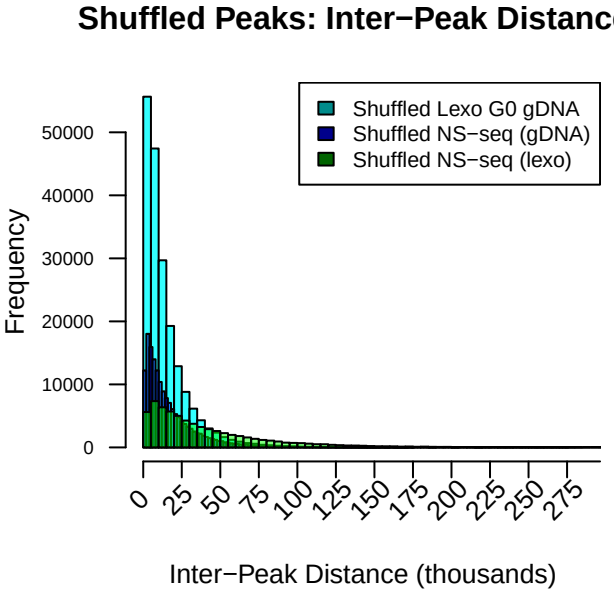

Peak Summit log10(−log10(qVal))
log10(−log10(qVal))
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)

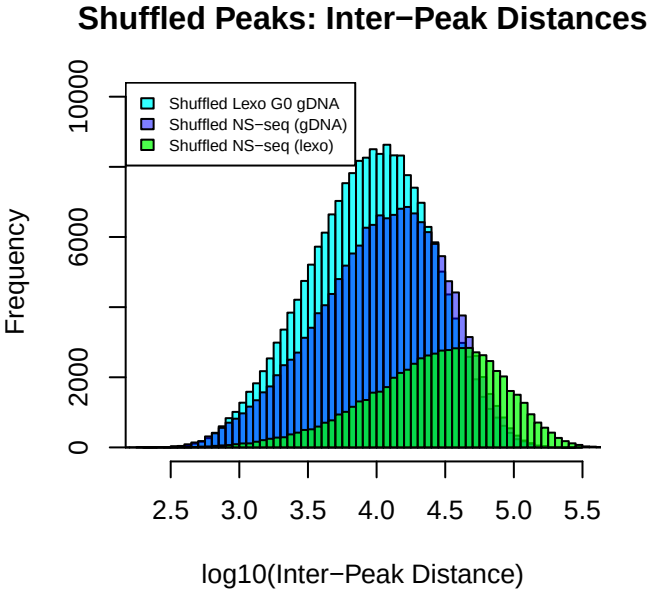

Inter−Peak Distances
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Frequency
140000
120000
100000
80000
60000
40000
20000
0
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
Inter−Peak Distance (thousands)

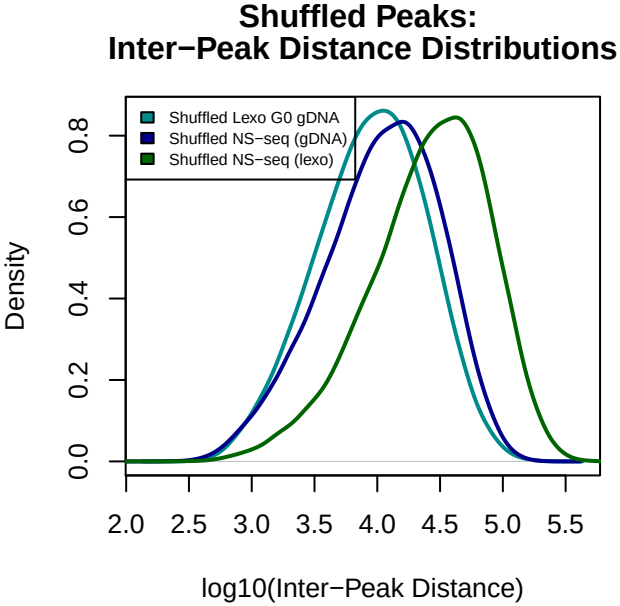

Inter−Peak Distances
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Frequency
8000
6000
4000
2000
0
2
3
4
5
6
log10(Inter−Peak Distance)

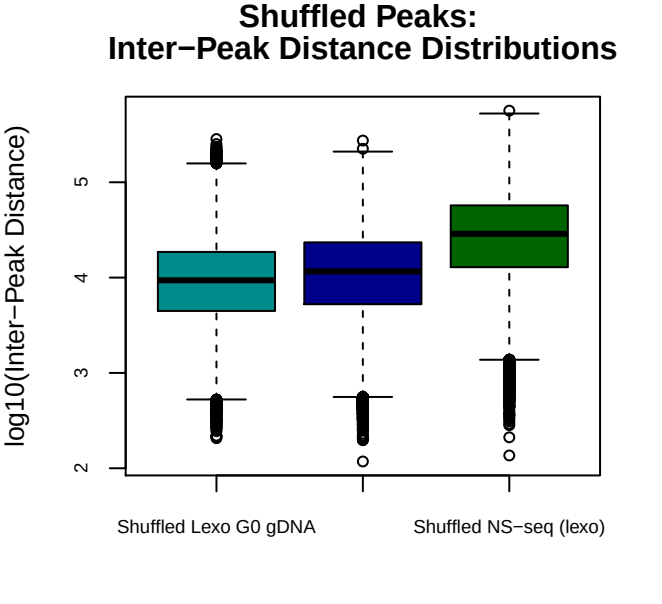

Inter−Peak Distance Distributions
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Density
0.8
0.6
0.4
0.2
0.0
2
3
4
5
6
log10(Inter−Peak Distance)

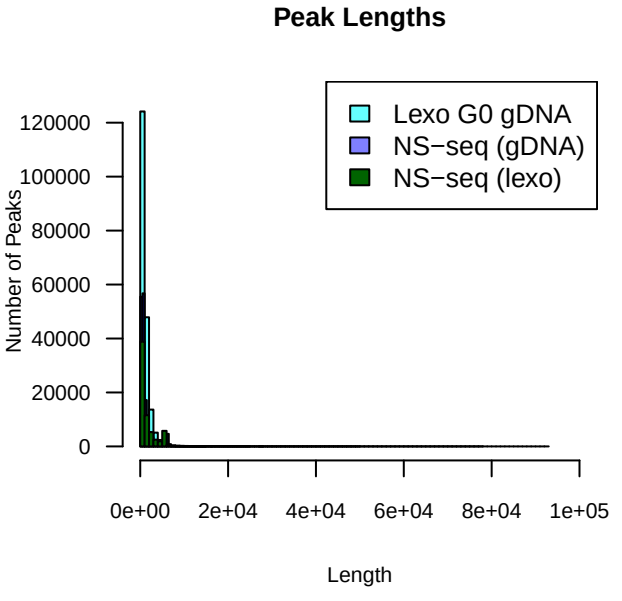

Inter−Peak Distance Distributions
log10(Inter−Peak Distance)
6
5
4
3
2
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)

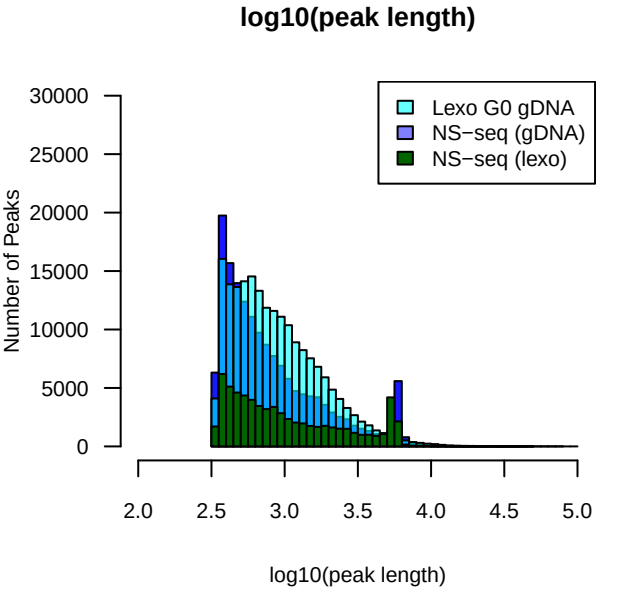

Shuffled Peaks: Inter−Peak Distances
Shuffled Lexo G0 gDNA
Shuffled NS−seq (gDNA)
Shuffled NS−seq (lexo)
Frequency
50000
40000
30000
20000
10000
0
0
25
50
75
100
125
150
175
200
225
250
275
Inter−Peak Distance (thousands)

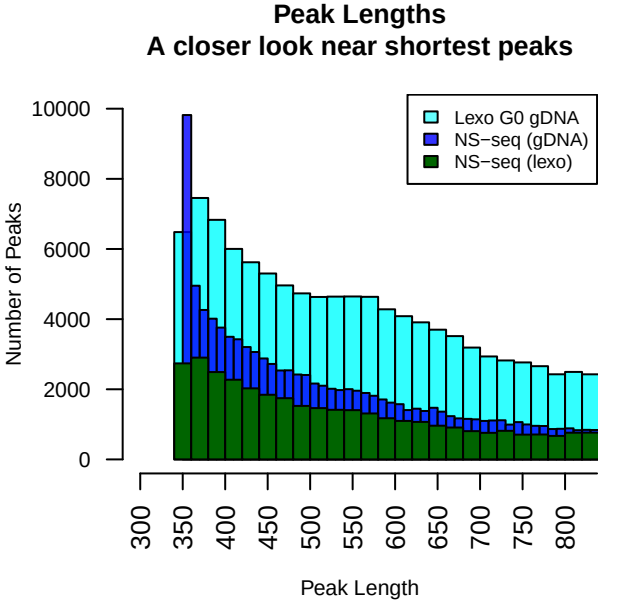

Shuffled Peaks: Inter−Peak Distances
Shuffled Lexo G0 gDNA
Shuffled NS−seq (gDNA)
Shuffled NS−seq (lexo)
Frequency
10000
6000
2000
0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
log10(Inter−Peak Distance)

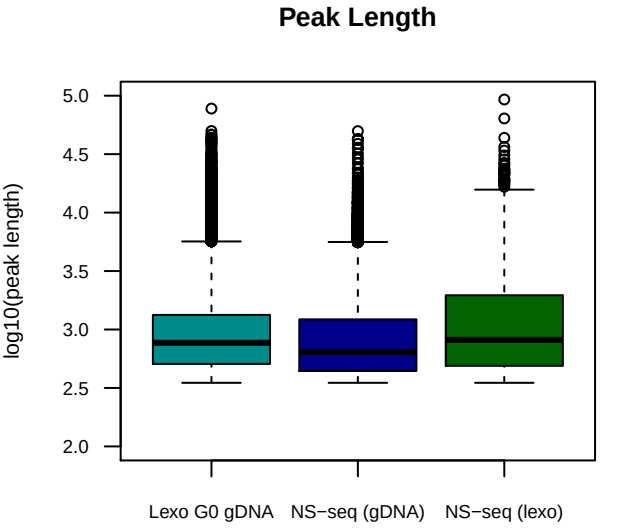

Shuffled Peaks:
Inter−Peak Distance Distributions
Shuffled Lexo G0 gDNA
Shuffled NS−seq (gDNA)
Shuffled NS−seq (lexo)
Density
0.8
0.6
0.4
0.2
0.0
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
log10(Inter−Peak Distance)

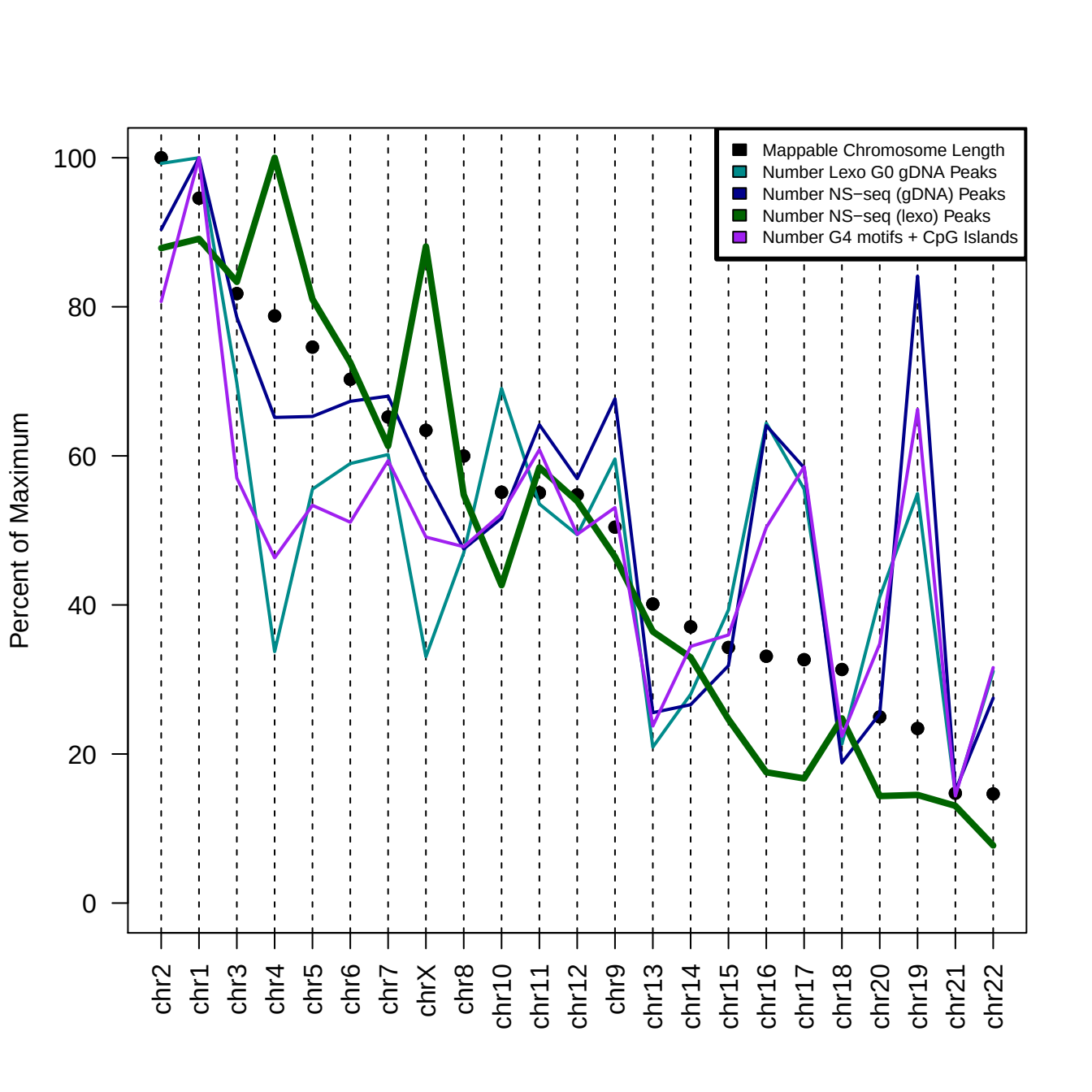

Shuffled Peaks:
Inter−Peak Distance Distributions
log10(Inter−Peak Distance)
5
4
3
2
Shuffled Lexo G0 gDNA
Shuffled NS−seq (lexo)

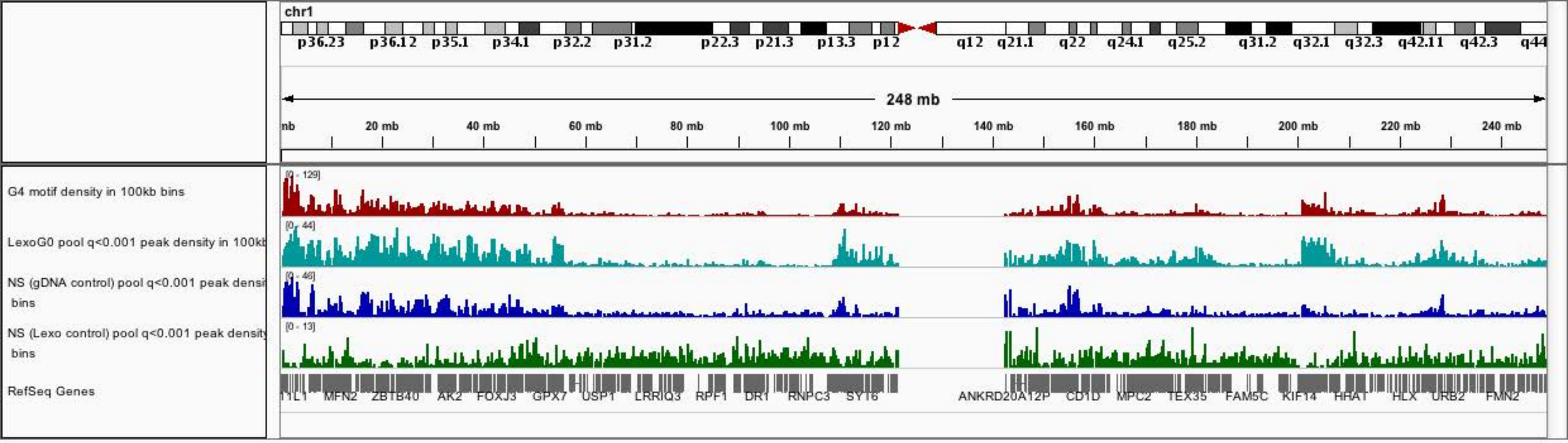

Peak Lengths
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
120000
100000
80000
60000
40000
20000
0
0e+00
2e+04
4e+04
6e+04
8e+04
1e+05
Length

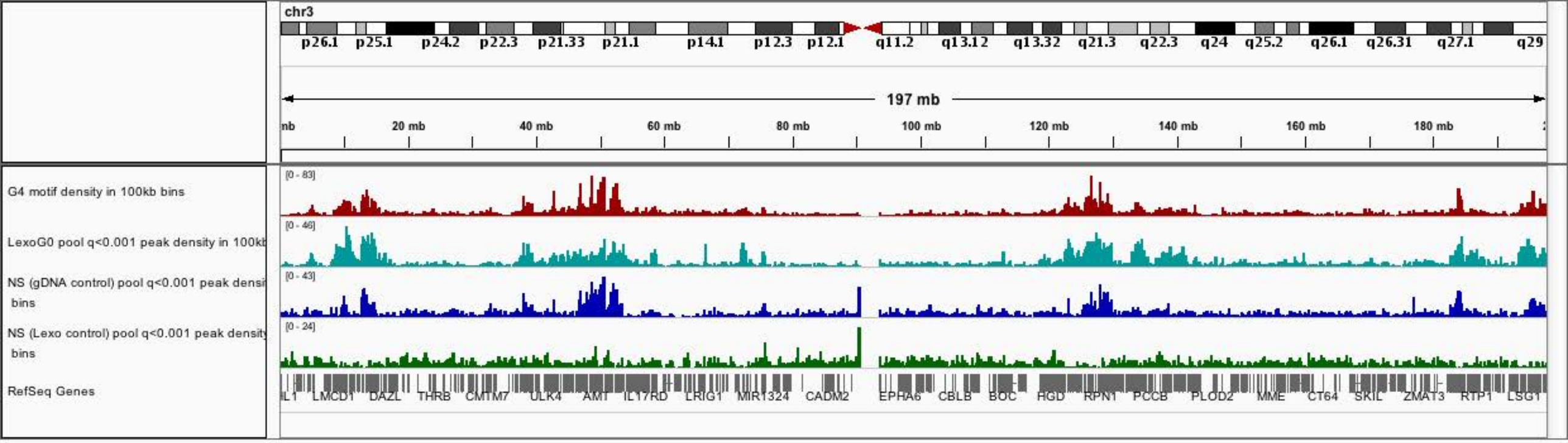

log10(peak length)
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
30000
25000
20000
15000
10000
5000
0
2.0
2.5
3.0
3.5
4.0
4.5
5.0
log10(peak length)

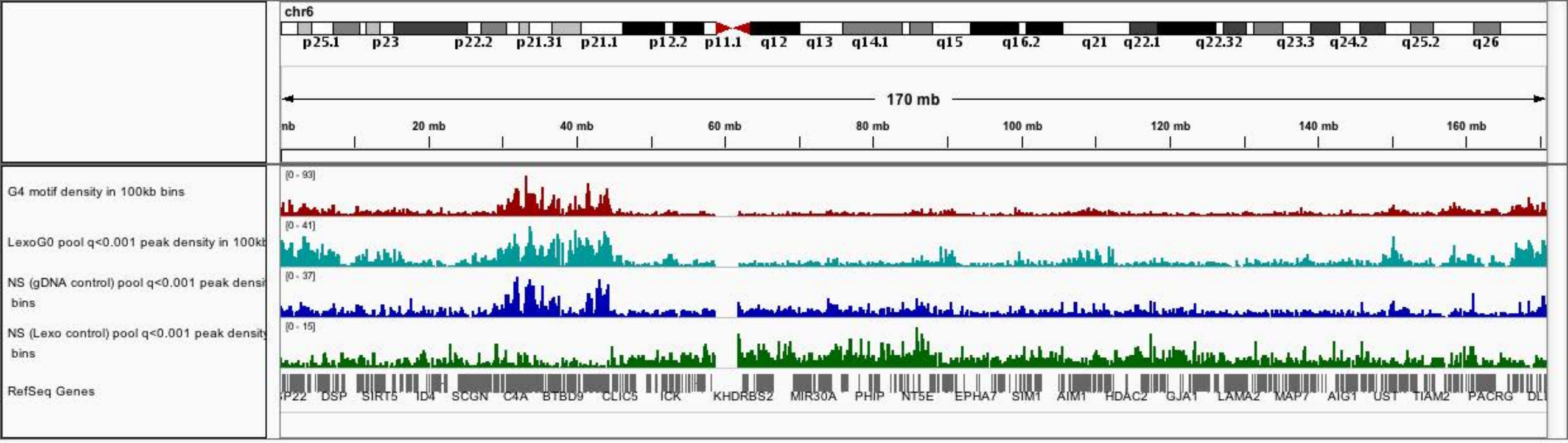

Peak Lengths
A closer look near shortest peaks
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Number of Peaks
10000
8000
6000
4000
2000
0
300
350
400
450
500
550
600
650
700
750
800
Peak Length

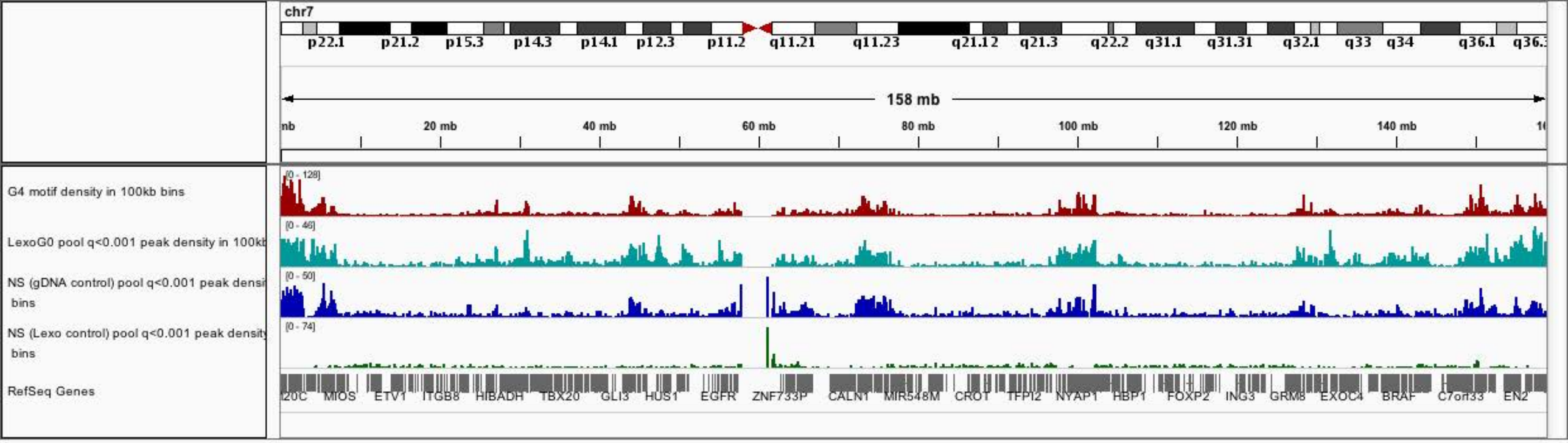

Peak Length
log10(peak length)
5.0
4.5
4.0
3.5
3.0
2.5
2.0
Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)

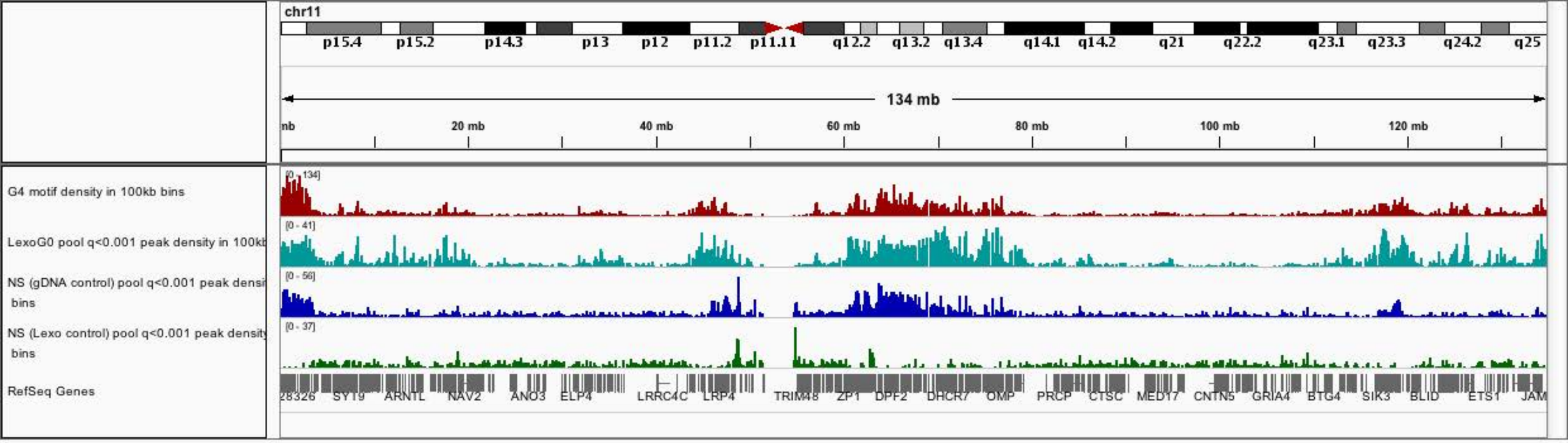

Mappable Chromosome Length
Number Lexo G0 gDNA Peaks
Number NS−seq (gDNA) Peaks
Number NS−seq (lexo) Peaks
Number G4 motifs + CpG Islands
Percent of Maximum
100
80
60
40
20
0
chr2
chr1
chr3
chr4
chr5
chr6
chr7
chrX
chr8
chr10
chr11
chr12
chr9
chr13
chr14
chr15
chr16
chr17
chr18
chr20
chr19
chr21
chr22

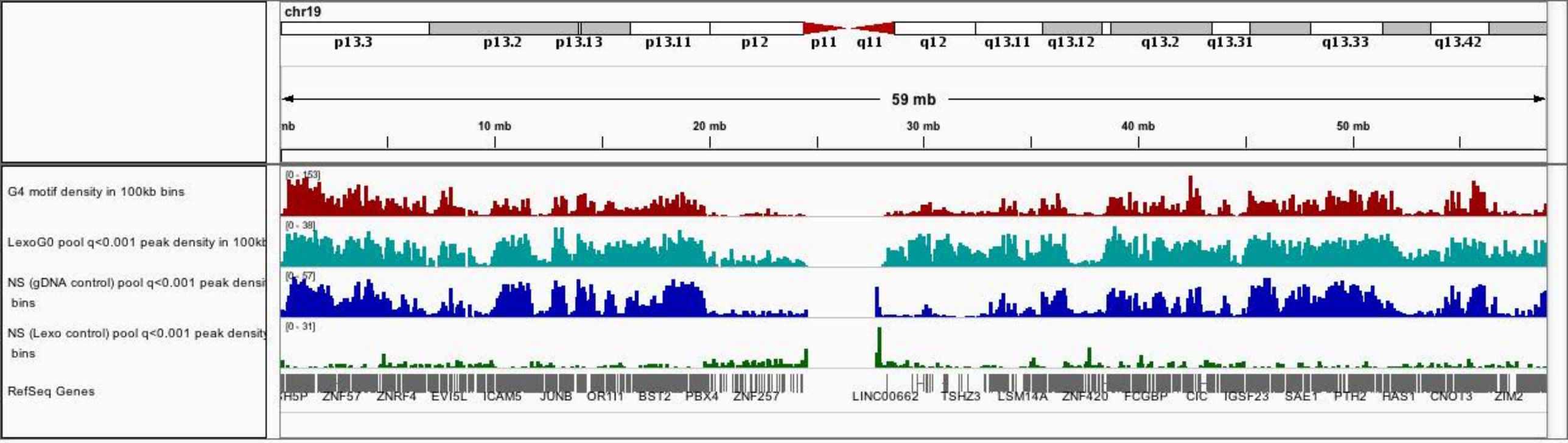

chr1
p36.23 p36.12 p35.1 p34.1 p32.2 p31.2 p22.3 p21.3 p13.3 p12 q12 q21.1 q22 q24.1 q25.2 q31.2 q32.1 q32.3 q42.11 q42.3 q44
248 mb
nb 20 mb 40 mb 60 mb 80 mb 100 mb 120 mb 140 mb 160 mb 180 mb 200 mb 220 mb 240 mb
G4 motif density in 100kb bins
[0 - 129]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 44]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 46]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 13]
RefSeq Genes
11L1 MFN2 ZBTB40 AK2 FOXJ3 GPX7 USP1 LRRIQ3 RPF1 DR1 RNPC3 SYT6 ANKRD20A12P CD1D MPC2 TEX35 FAM5C KIF14 HHAT HLX URB2 FMN2

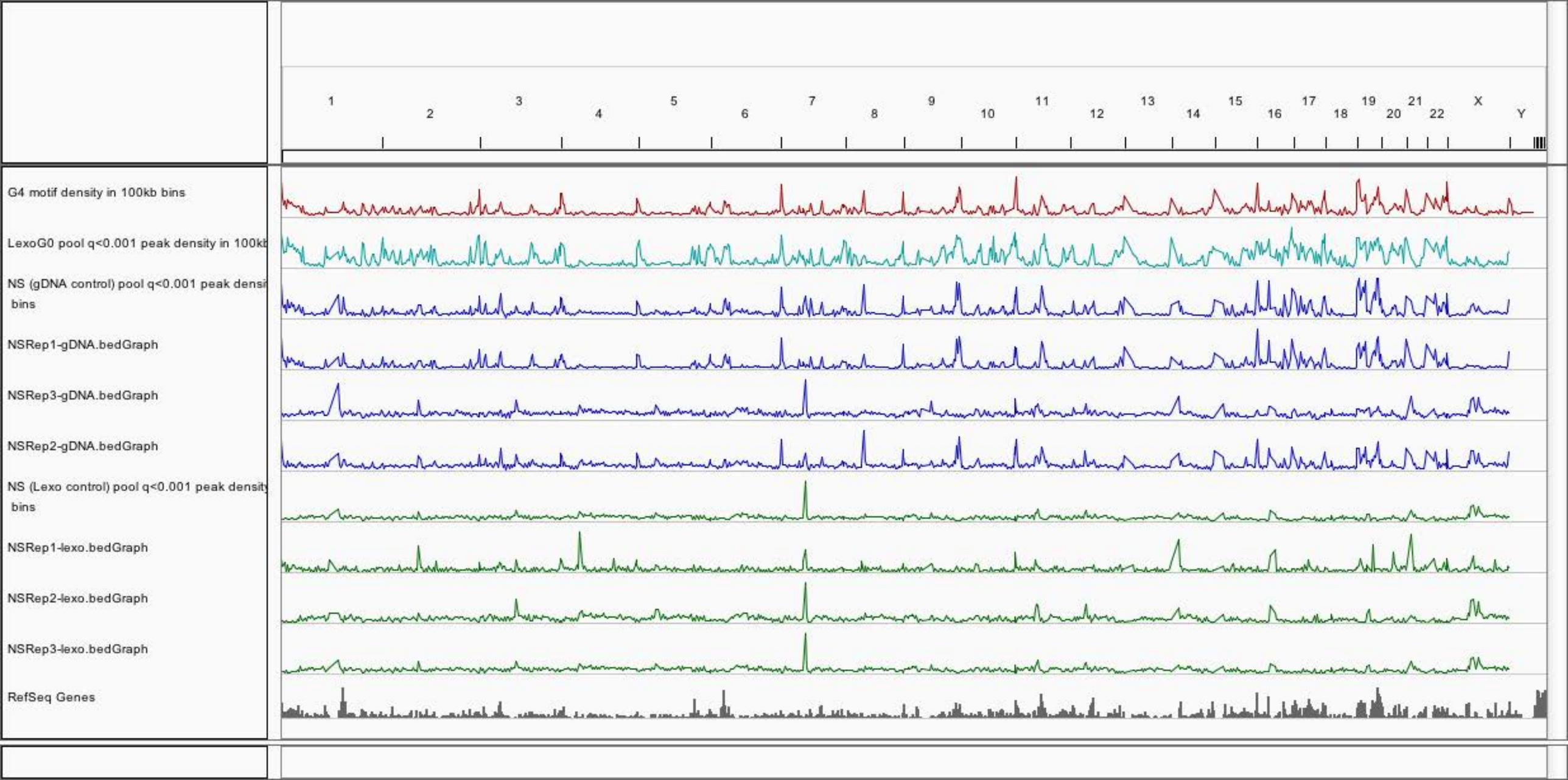

chr3
p26.1 p25.1 p24.2 p22.3 p21.33 p21.1 p14.1 p12.3 p12.1 q11.2 q13.12 q13.32 q21.3 q22.3 q24 q25.2 q26.1 q26.31 q27.1 q29
197 mb
nb 20 mb 40 mb 60 mb 80 mb 100 mb 120 mb 140 mb 160 mb 180 mb
G4 motif density in 100kb bins
[0 - 83]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 46]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 43]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 24]
RefSeq Genes
IL1 LMCD1 DAZL THRB CMTM7 ULK4 AMT IL17RD LRIG1 MIR1324 CADM2 EPHA6 CBLB BOC HGD RPN1 PCCB PLOD2 MME CT64 SKIL ZMAT3 RTP1 LSG1

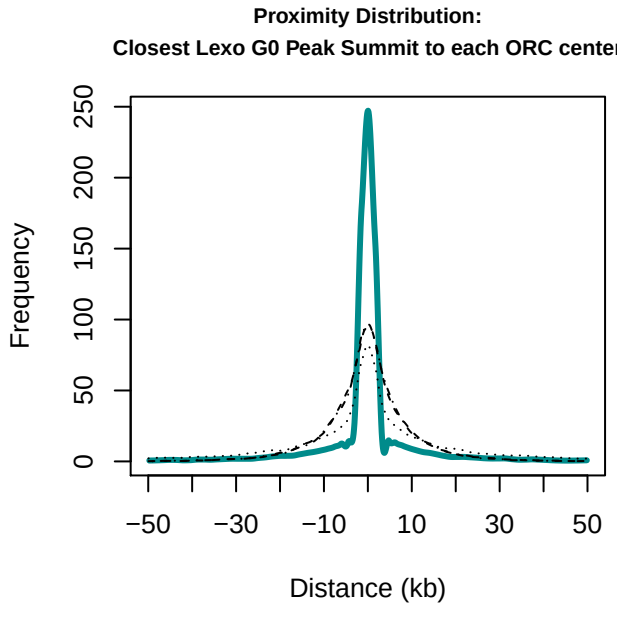

chr6
p25.1 p23 p22.2 p21.31 p21.1 p12.2 p11.1 q12 q13 q14.1 q15 q16.2 q21 q22.1 q22.32 q23.3 q24.2 q25.2 q26
170 mb
nb 20 mb 40 mb 60 mb 80 mb 100 mb 120 mb 140 mb 160 mb
G4 motif density in 100kb bins
[0 - 93]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 41]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 37]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 15]
RefSeq Genes
P22 DSP SIRT5 ID4 SCGN C4A BTBD9 CLIC5 ICK KHDRBS2 MIR30A PHIP NT5E EPHA7 SIM1 AIM1 HDAC2 GJA1 LAMA2 MAP7 AIG1 UST TIAM2 PACRG DL

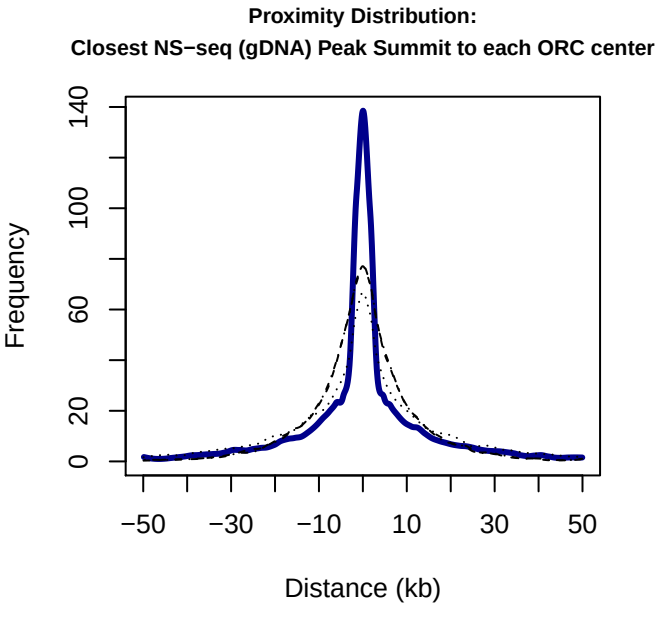

chr7
p22.1
p21.2
p15.3
p14.3
p14.1
p12.3
p11.2
q11.21
q11.23
q21.12
q21.3
q22.2
q31.1
q31.31
q32.1
q33
q34
q36.1
158 mb
20 mb
40 mb
60 mb
80 mb
100 mb
120 mb
140 mb
G4 motif density in 100kb bins
[0 - 128]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 46]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 50]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 74]
RefSeq Genes
I20C
MIOS
ETV1
ITGB8
HIBADH
TBX20
GLI3
HUS1
EGFR
ZNF733P
CALN1
MIR548M
CROT
TFPI2
NYAP1
HBP1
FOXP2
ING3
GRM8
EXOC4
BRAF
C7orf33
EN2

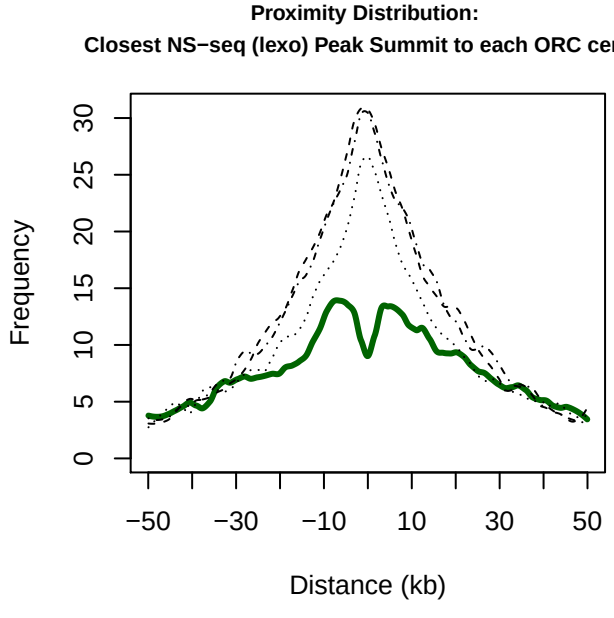

chr11
p15.4
p15.2
p14.3
p13
p12
p11.2
p11.11
q12.2
q13.2
q13.4
q14.1
q14.2
q21
q22.2
q23.1
q23.3
q24.2
q25
134 mb
nb
20 mb
40 mb
60 mb
80 mb
100 mb
120 mb
G4 motif density in 100kb bins
[0 - 134]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 41]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 56]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 37]
RefSeq Genes
28326
SYT9
ARNTL
NAV2
ANO3
ELP4
LRRC4C
LRP4
TRIM48
ZP1
DPF2
DHCR7
OMP
PRCP
CTSC
MED17
CNTN5
GRIA4
BTG4
SIK3
BLID
ETS1
JAM

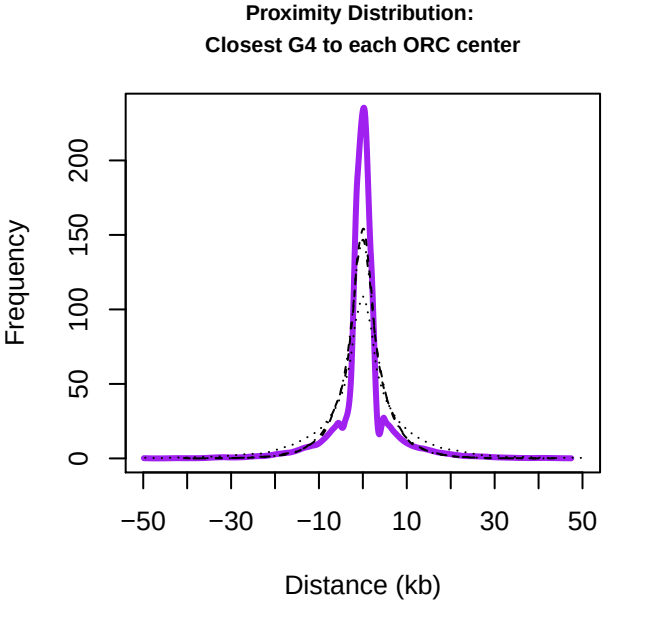

chr19
p13.3
p13.2
p13.13
p13.11
p12
p11
q11
q12
q13.11
q13.12
q13.2
q13.31
q13.33
q13.42
59 mb
10 mb
20 mb
30 mb
40 mb
50 mb
G4 motif density in 100kb bins
[0 - 153]
LexoG0 pool q<0.001 peak density in 100kb
[0 - 38]
NS (gDNA control) pool q<0.001 peak densi
bins
[0 - 57]
NS (Lexo control) pool q<0.001 peak density
bins
[0 - 31]
RefSeq Genes
ZNF57
ZNRF4
EVI5L
ICAM5
JUNB
OR1I1
BST2
PBX4
ZNF257
LINC00662
TSHZ3
LSM14A
ZNF420
FCGBP
CIC
IGSF23
SAE1
PTH2
RAS1
CNOT3
ZIM2

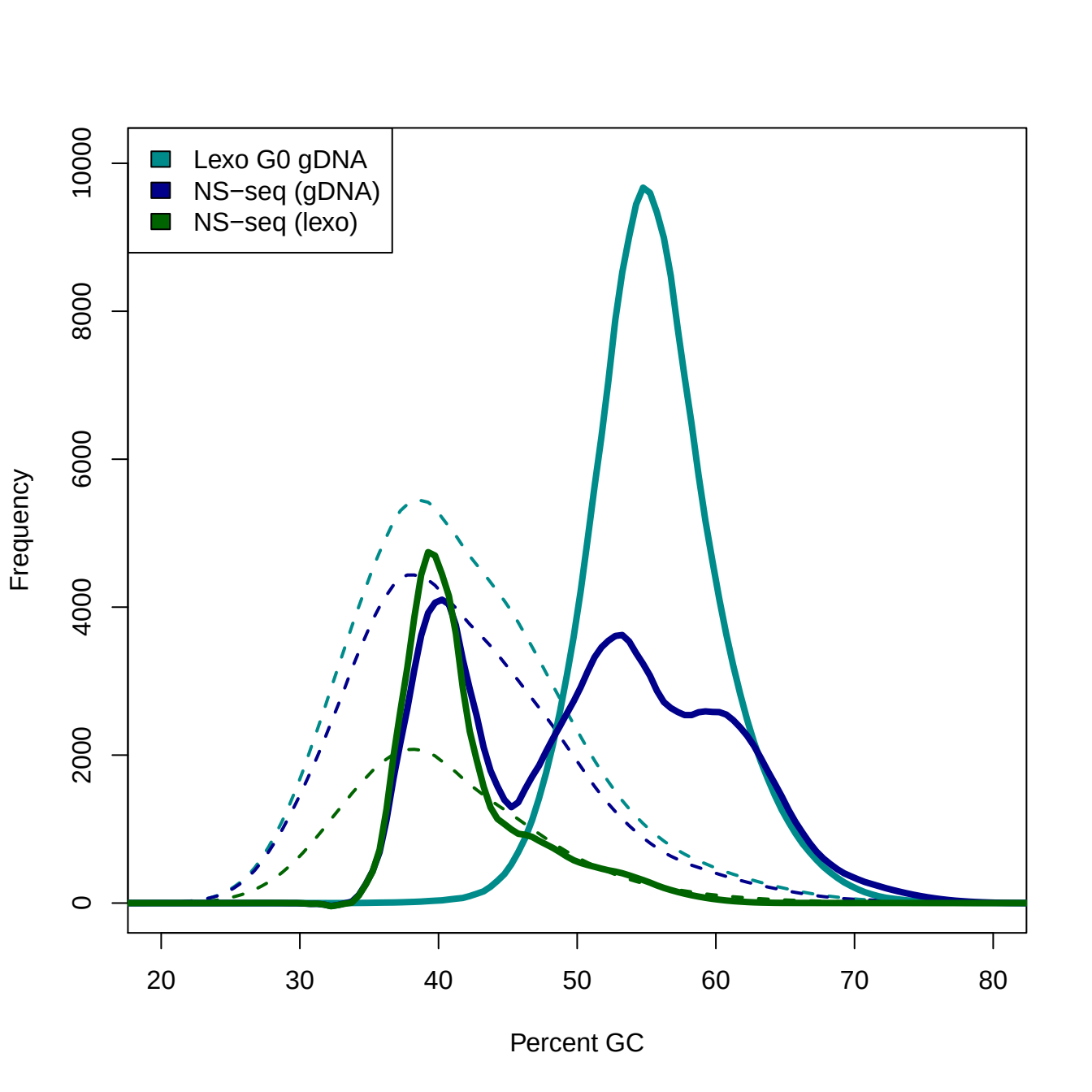

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
G4 motif density in 100kb bins
LexoG0 pool q<0.001 peak density in 100kb
NS (gDNA control) pool q<0.001 peak densi
bins
NSRep1-gDNA.bedGraph
NSRep3-gDNA.bedGraph
NSRep2-gDNA.bedGraph
NS (Lexo control) pool q<0.001 peak density
bins
NSRep1-lexo.bedGraph
NSRep2-lexo.bedGraph
NSRep3-lexo.bedGraph
RefSeq Genes

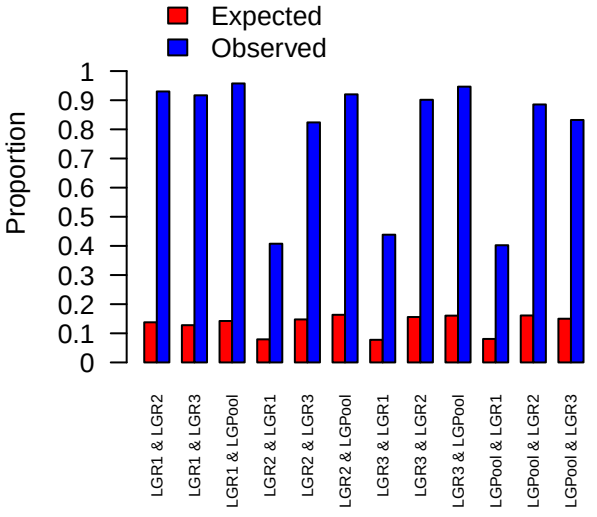

Proximity Distribution:
Closest Lexo G0 Peak Summit to each ORC center
Frequency
0
50
100
150
200
250
−50
−30
−10
10
30
50
Distance (kb)

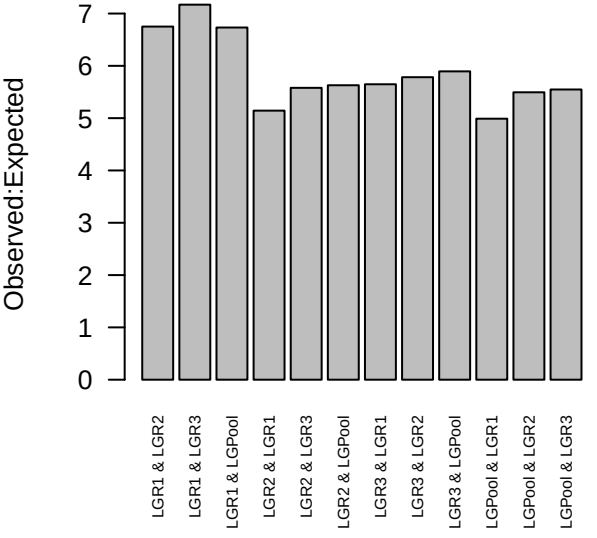

Proximity Distribution:
Closest NS−seq (gDNA) Peak Summit to each ORC center
Frequency
0
20
60
100
140
−50
−30
−10
10
30
50
Distance (kb)

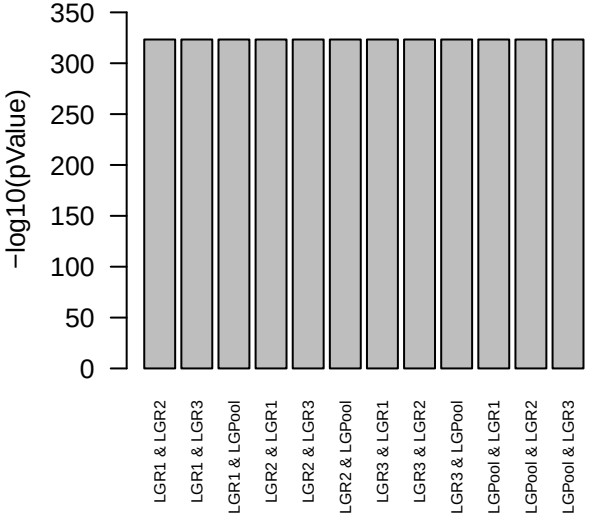

Proximity Distribution:
Closest NS−seq (lexo) Peak Summit to each ORC center
Frequency
0
5
10
15
20
25
30
−50
−30
−10
10
30
50
Distance (kb)

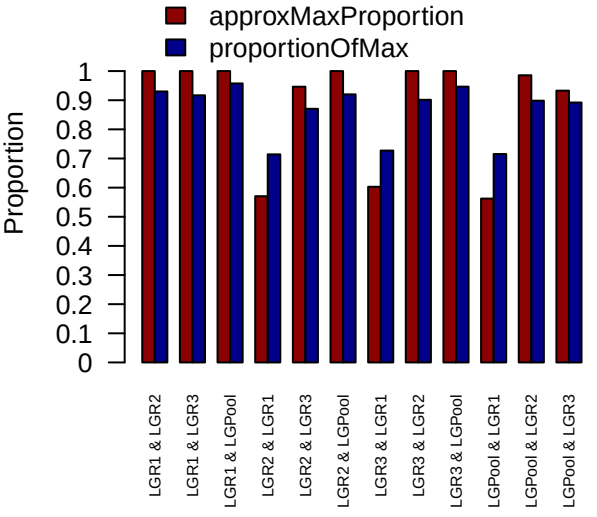

Proximity Distribution:
Closest G4 to each ORC center
Frequency
0
50
100
150
200
−50
−30
−10
10
30
50
Distance (kb)

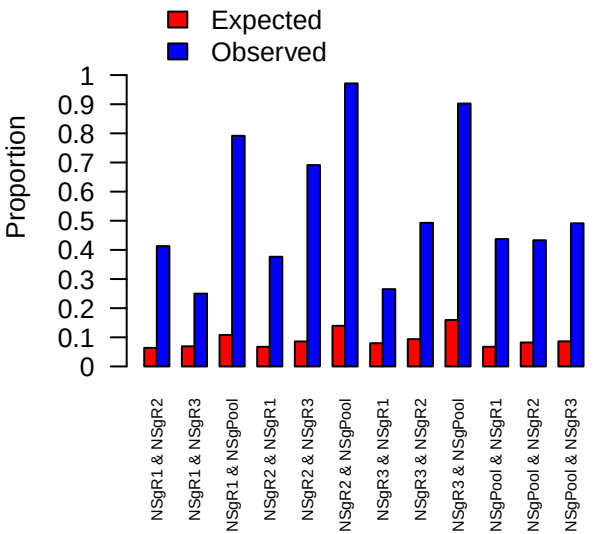

Lexo G0 gDNA
NS−seq (gDNA)
NS−seq (lexo)
Frequency
0
2000
4000
6000
8000
10000
20
30
40
50
60
70
80
Percent GC

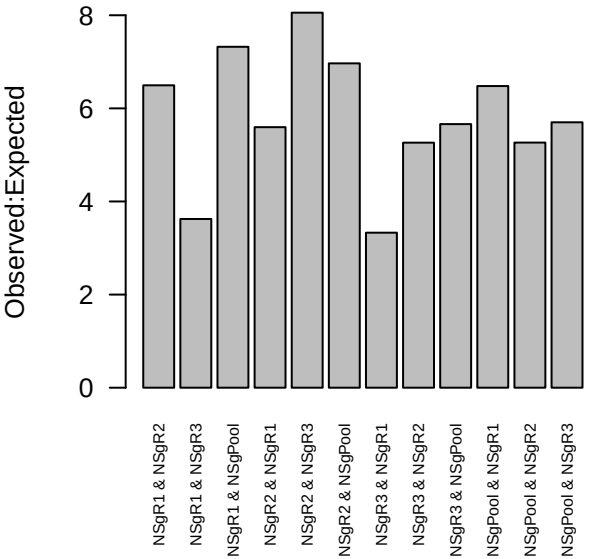

Expected
Observed
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
LGR1 & LGR2
LGR1 & LGR3
LGR1 & LGPool
LGR2 & LGR1
LGR2 & LGR3
LGR2 & LGPool
LGR3 & LGR1
LGR3 & LGR2
LGR3 & LGPool
LGPool & LGR1
LGPool & LGR2
LGPool & LGR3

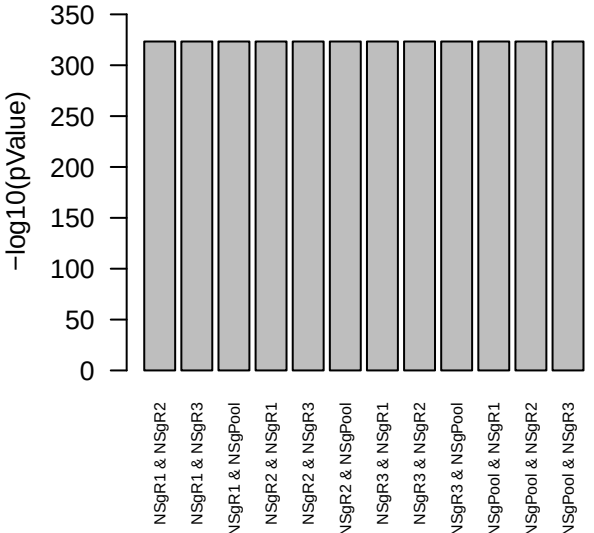

Observed:Expected
0
1
2
3
4
5
6
7
LGR1 & LGR2
LGR1 & LGR3
LGR1 & LGPool
LGR2 & LGR1
LGR2 & LGR3
LGR2 & LGPool
LGR3 & LGR1
LGR3 & LGR2
LGR3 & LGPool
LGPool & LGR1
LGPool & LGR2
LGPool & LGR3

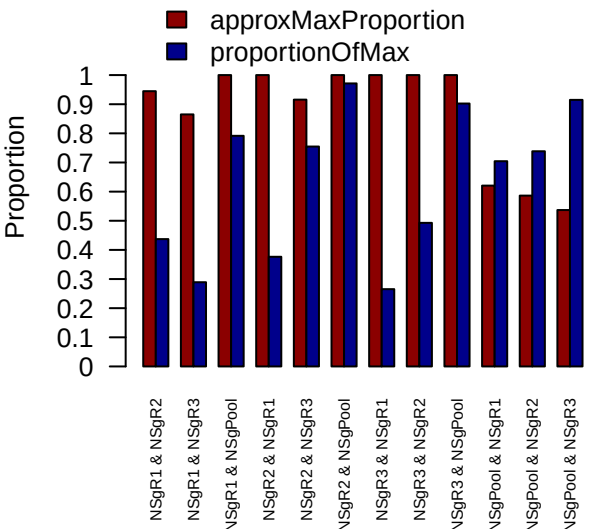

−log10(pValue)
0
50
100
150
200
250
300
350
LGR1 & LGR2
LGR1 & LGR3
LGR1 & LGPool
LGR2 & LGR1
LGR2 & LGR3
LGR2 & LGPool
LGR3 & LGR1
LGR3 & LGR2
LGR3 & LGPool
LGPool & LGR1
LGPool & LGR2
LGPool & LGR3

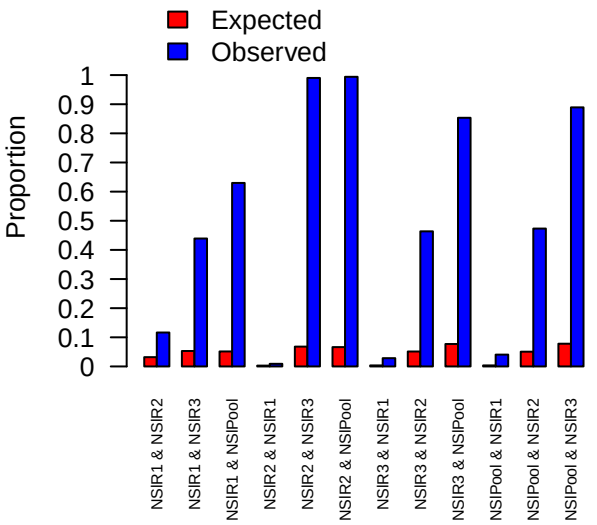

approxMaxProportion
proportionOfMax
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
LGR1 & LGR2
LGR1 & LGR3
LGR1 & LGPool
LGR2 & LGR1
LGR2 & LGR3
LGR2 & LGPool
LGR3 & LGR1
LGR3 & LGR2
LGR3 & LGPool
LGPool & LGR1
LGPool & LGR2
LGPool & LGR3

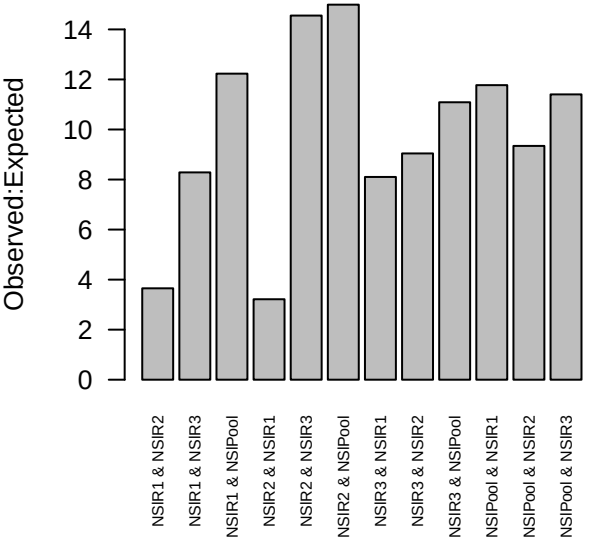

Expected
Observed
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
NSgR1 & NSgR2
NSgR1 & NSgR3
NSgR1 & NSgPool
NSgR2 & NSgR1
NSgR2 & NSgR3
NSgR2 & NSgPool
NSgR3 & NSgR1
NSgR3 & NSgR2
NSgR3 & NSgPool
NSgPool & NSgR1
NSgPool & NSgR2
NSgPool & NSgR3

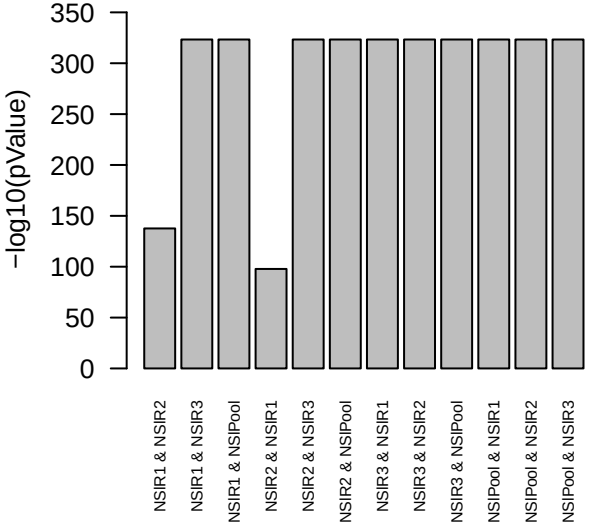

Observed:Expected
0
2
4
6
8
NSgR1 & NSgR2
NSgR1 & NSgR3
NSgR1 & NSgPool
NSgR2 & NSgR1
NSgR2 & NSgR3
NSgR2 & NSgPool
NSgR3 & NSgR1
NSgR3 & NSgR2
NSgR3 & NSgPool
NSgPool & NSgR1
NSgPool & NSgR2
NSgPool & NSgR3

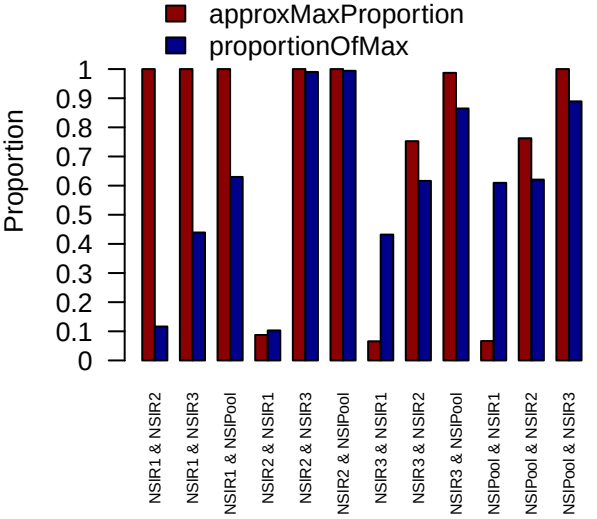

−log10(pValue)
0
50
100
150
200
250
300
350
NSgR1 & NSgR2
NSgR1 & NSgR3
NSgR1 & NSgPool
NSgR2 & NSgR1
NSgR2 & NSgR3
NSgR2 & NSgPool
NSgR3 & NSgR1
NSgR3 & NSgR2
NSgR3 & NSgPool
NSgPool & NSgR1
NSgPool & NSgR2
NSgPool & NSgR3

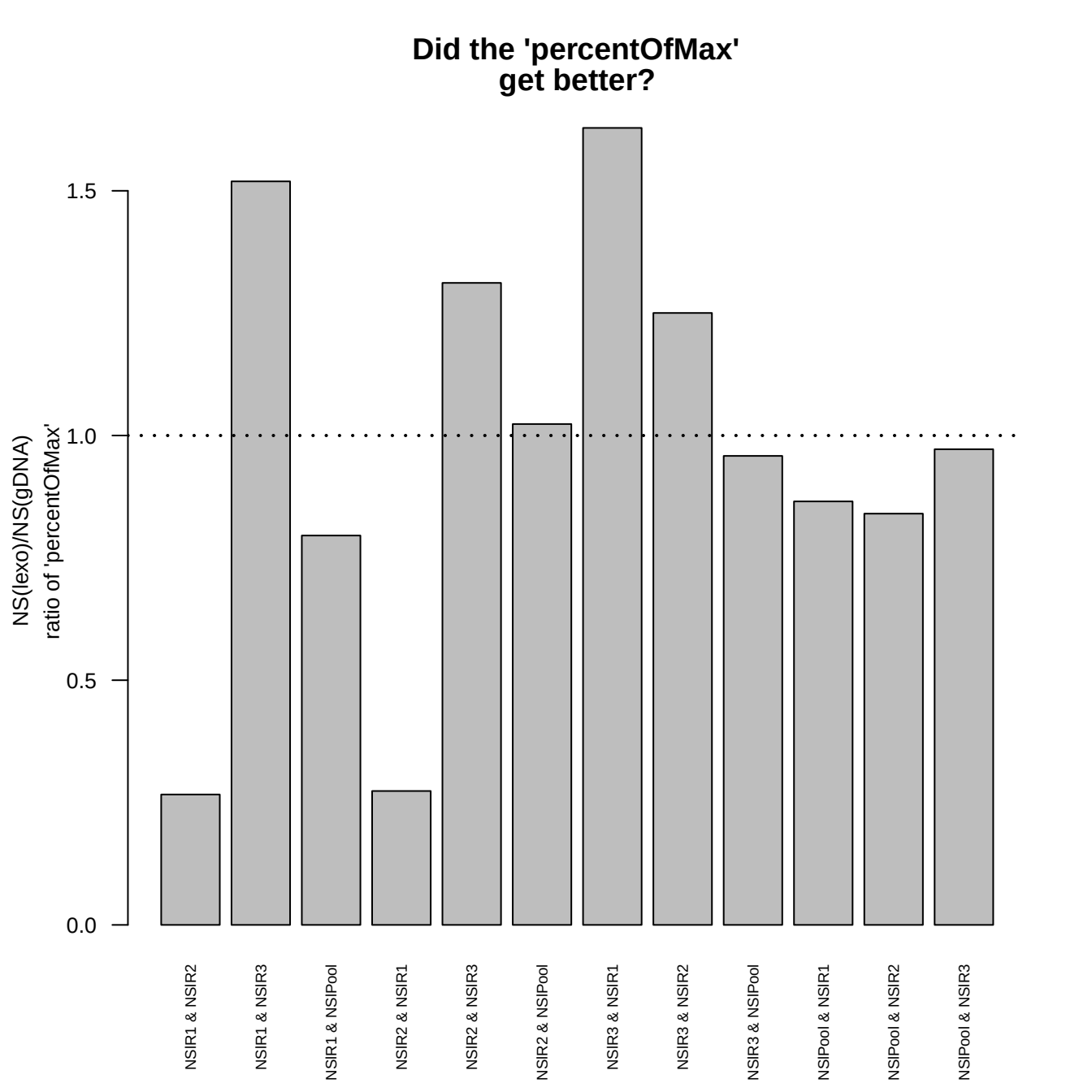

approxMaxProportion
proportionOfMax
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
NSgR1 & NSgR2
NSgR1 & NSgR3
NSgR1 & NSgPool
NSgR2 & NSgR1
NSgR2 & NSgR3
NSgR2 & NSgPool
NSgR3 & NSgR1
NSgR3 & NSgR2
NSgR3 & NSgPool
NSgPool & NSgR1
NSgPool & NSgR2
NSgPool & NSgR3

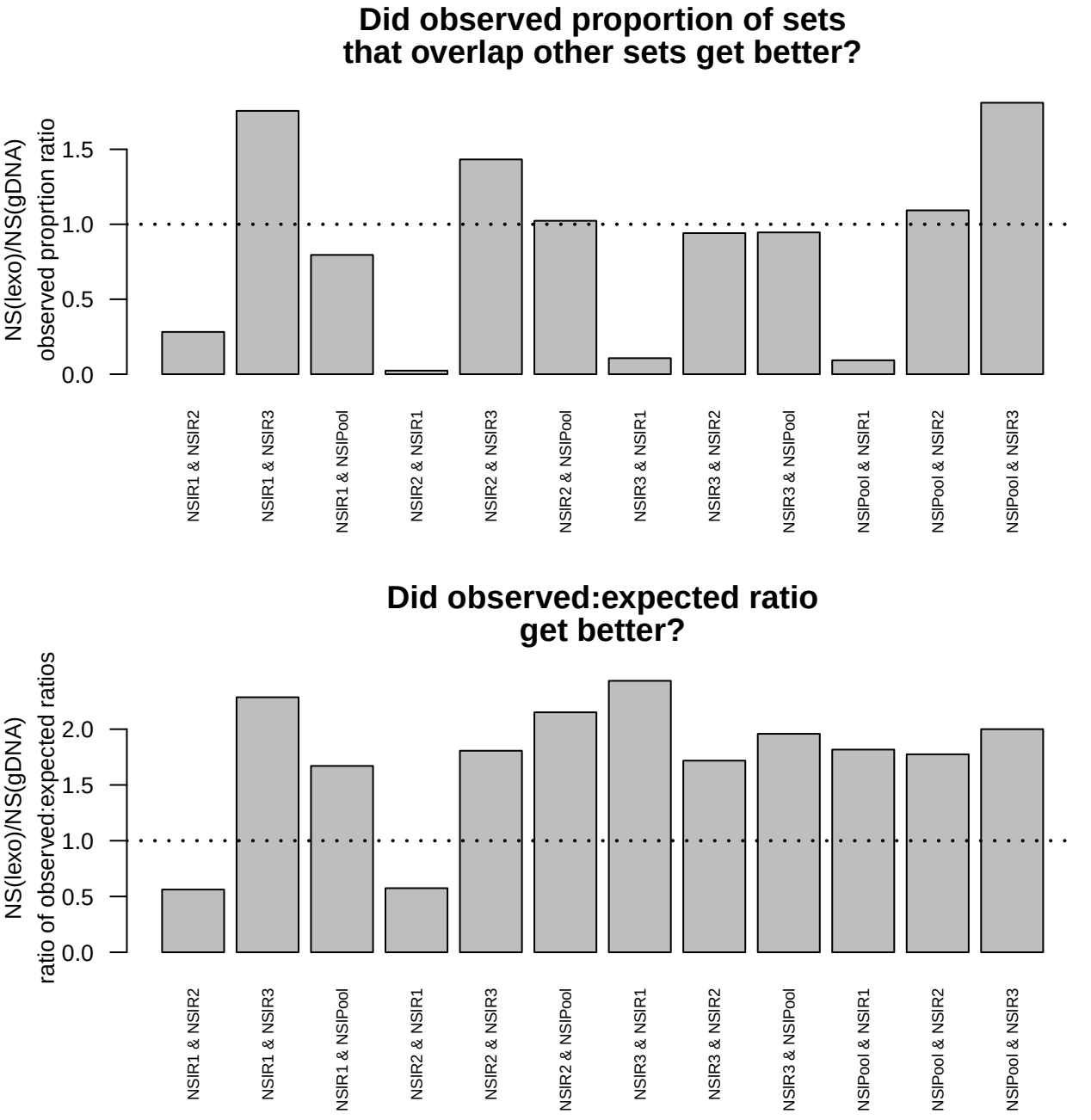

Expected
Observed
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3


Observed:Expected
0
2
4
6
8
10
12
14
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3


−log10(pValue)
0
50
100
150
200
250
300
350
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3


approxMaxProportion
proportionOfMax
Proportion
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3


Did the 'percentOfMax'
get better?
NS(lexo)/NS(gDNA)
ratio of 'percentOfMax'
0.0
0.5
1.0
1.5
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3


Did observed proportion of sets
that overlap other sets get better?
NS(lexo)/NS(gDNA)
observed proportion ratio
0.0
0.5
1.0
1.5
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3
Did observed:expected ratio
get better?
NS(lexo)/NS(gDNA)
ratio of observed:expected ratios
0.0
0.5
1.0
1.5
2.0
NSIR1 & NSIR2
NSIR1 & NSIR3
NSIR1 & NSIPool
NSIR2 & NSIR1
NSIR2 & NSIR3
NSIR2 & NSIPool
NSIR3 & NSIR1
NSIR3 & NSIR2
NSIR3 & NSIPool
NSIPool & NSIR1
NSIPool & NSIR2
NSIPool & NSIR3

Number of Reads In Each Data Set - numReads

Sample Name	Date	Platform	In use for paper?	Read Length	Total numReads	numMappableReads*	percent Mappable	numReads mapped to one location	Of Mapped Reads Percent Mapped To One Location	num of all mappable reads that passed MACS2 redundant read filter keeping only 1 read at each redundant location	percent of total mappable reads that was redundant and eliminated (redundant rate)	num Mappable reads with MAPQ >= 30	percent of Mappable Reads that have MAPQ >= 30	num mapq >=30 and non-redundant (when MACS2 keeps only 1)	redundant rate of mappable reads with MAPQ>=30	Fraction of All Mappable Reads in Telomeres (FRiT_allMap)	Fraction of all mappable redundant-filtered reads in Telomeres (FRiT_allMapNonRdt)	Fraction of Mappable Reads w/ MAPQ >= 30 in Telomeres (FRiT_allMap_Q30)	Fraction of mappable redundant-filtered reads with MAPQ >= 30 in Telomeres (FRiT_allMapNonRdt_
NS Rep1 Pilot	5/19/2011	GAIIx	No	42	26587570	12238556	46.0311190530011	NA	NA	NA	NA								
NS Rep1	9/29/2011	HiSeq2000	Yes	50	128247879	57397008	44.7547424936361	29297485	51.0435753027405	45124630	21.381564000688	37183014	64.7821468324621	30871035	16.9754366873003				
NS Rep2	9/29/2011	HiSeq2000	Yes	50	139320824	89354586	64.135843755848	47561217	53.2275053011829	71264584	20.2451858486592	57940694	64.8435593445646	51096712	11.8120469872177				
NS Rep3	9/29/2011	HiSeq2000	Yes	50	136227515	96762425	71.0300154854913	33746889	34.8760265154578	67753545	29.9794884222879	49371654	51.0235806926087	40553346	17.8610746968291				
G0 gDNA	9/29/2011	HiSeq2000	Yes	50	193565007	181911420	93.97949702758	137010993	75.317422622505	165565467	8.98566621051059	154890949	85.1463580461304	144804056	6.5122546314827	0.00003062479529			
LexoG0 Rep1	1/31/2013	HiSeq2000	Yes	50	162102273	115150124	71.035477707336	75075607	65.1980253186701	57272428	50.262816911947	86965603	75.5236729054673	41788662	51.9480569806432				
LexoG0 Rep2	5/23/2013	HiSeq2000	Yes	50	196524435	174625028	88.8566493016504	118958676	68.1223518543974	132316521	24.2282034165229	134964477	77.2881634126352	105892838	21.5402153560748				
LexoG0 Rep3	5/23/2013	HiSeq2000	Yes	50	174860103	153657189	87.8743557642763	104359761	67.9172654915612	115535011	24.8098889795518	117645145	76.5633848735838	92494676	21.3782464206236				
Pooled NS reps	9/29/2011	HiSeq2000	Yes	50	403796218	243514019	60.3061663643417	110605591	45.4206256601596	184142759	24.3810439513135	144495362	59.3375948511613	122521093	15.2075946908247	0.00192485016642			
Pooled Lexo Reps	1/31/2013 and 5/23/2013	HiSeq2000	Yes	50	533486811	443432341	83.1196445454394	298394044	67.291899216706	305123960	31.1904135562363	339575225	76.5788134068462	240176176	29.2715845215151	0.00037567399713			

Reads were mapped with bowtie2 for these quotes: bowtie2 -p 8 -t --very-sensitive -N 1 -x hg19Index -U file.fq -S outputPrefix

*Rep1 Pilot (GAIIx) quote is from hg18 mapping. All others are from hg19 mapping.

Note: For pooled datasets, individual replicates were processed individually and merged rather than merging raw data (fastq) and processing.

This ensures that reads from different replicates are not treated as redundant if they occur at the same position. It avoids eliminating real data.

Note: For our first paper, we have used all mappable reads in our peak calling, signal track correlations, etc. Thus, the number of reads that map to just one position is just FYI.

It gives an idea of how many reads we would have if we went that route (which is different than the MAPQ>=30 route).

Groups that used reads that mapped just once: Aladjem, ...

numPeaks - MCF7 NS-seq - MACS2

Sample	Control	Read Processing (mapq, filterdup, etc)	Cut Off	Number of Peaks	Parameter Notes	peak length filtering	chrY,chrM peaks removed?	Other filtering?	Date	MACS2 Version	Purpose	If down(up) sampling, was treat or control down(up) sampled?	num treatment Reads	num Control Reads	num reads down (up) sampled from	percent of reads sampled	Percent Difference: (numTreat - numControl) /numTreat	Seed for reproducing downsampling	Command
NSRep1	gDNA	all mappable	p<1e-3	150304		all lengths	No	No	4/14/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969		
NSRep1	gDNA	all mappable	p<1e-3	180477		all lengths	No	No	7/2/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep1	gDNA	all mappable	p<1e-3	46142		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
NSRep1	gDNA	all mappable	p<1e-5	84940		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep1	gDNA	all mappable	p<1e-5	22606		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
NSRep1	gDNA	all mappable	q<0.05	254006		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep1	gDNA	all mappable	q<0.01	169801		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep1	gDNA	all mappable	q<0.001	100888		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	45124630	45124604	165565467	27.2548405278258	0.000057618200969	6583	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep1	gDNA	MAPQ>=30	p<1e-5	73351		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	30871035	30871010	144804056	21.3191611152108	0.00008098205972	8134	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep1	gDNA	MAPQ>=30	q<0.05	161144		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	30871035	30871010	144804056	21.3191611152108	0.00008098205972	6583	
NSRep1	gDNA	MAPQ>=30	q<0.01	122976		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	30871035	30871010	144804056	21.3191611152108	0.00008098205972	6583	
NSRep1	gDNA	MAPQ>=30	q<0.001	84947		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	30871035	30871010	144804056	21.3191611152108	0.00008098205972	6583	
NSRep2	gDNA	all mappable	p<1e-3	107491		all lengths	No	No	4/14/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep2	gDNA	all mappable	p<1e-3	113523		all lengths	No	No	7/2/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep2	gDNA	all mappable	p<1e-3	57417		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
NSRep2	gDNA	all mappable	p<1e-5	90126		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	all mappable	p<1e-5	43877		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
NSRep2	gDNA	all mappable	q<0.05	122741		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	all mappable	q<0.01	107906		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	all mappable	q<0.001	95929		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	71264584	71264563	165565467	43.0431322976306	0.000029467652544	22353	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	MAPQ>=30	p<1e-5	35119		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	51096712	51096689	144804056	35.2867802266533	0.000045012681051	11600	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed 11600 -g hs -n $outputPrefix -p 1e-5 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep2	gDNA	MAPQ>=30	q<0.05	76760		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	51096712	51096689	144804056	35.2867802266533	0.000045012681051	11600	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_05/$outputPrefix05 -q $CUTOFF05 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	MAPQ>=30	q<0.01	54410		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	51096712	51096689	144804056	35.2867802266533	0.000045012681051	11600	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_01/$outputPrefix01 -q $CUTOFF01 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	gDNA	MAPQ>=30	q<0.001	76760		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	51096712	51096689	144804056	35.2867802266533	0.000045012681051	11600	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_001/$outputPrefix001 -q $CUTOFF001 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	all mappable	p<1e-3	104997		all lengths	No	No	4/14/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep3	gDNA	all mappable	p<1e-3	116087		all lengths	No	No	7/2/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep3	gDNA	all mappable	p<1e-3	54158		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
NSRep3	gDNA	all mappable	p<1e-5	78912		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	all mappable	p<1e-5	43957		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
NSRep3	gDNA	all mappable	q<0.05	135207		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	all mappable	q<0.01	106673		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	all mappable	q<0.001	88200		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	67753545	67753521	165565467	40.9224956312901	0.000035422500771	19924	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	MAPQ>=30	p<1e-5	53279		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	12741	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep3	gDNA	MAPQ>=30	q<0.05	112615		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	12741	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_05/$outputPrefix05 -q $CUTOFF05 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	MAPQ>=30	q<0.01	82326		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	12741	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_01/$outputPrefix01 -q $CUTOFF01 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	gDNA	MAPQ>=30	q<0.001	60500		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	12741	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_001/$outputPrefix001 -q $CUTOFF001 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSpooled	gDNA	all mappable	p<1e-3	210072		all lengths	No	No	4/15/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	NS downsampled	165565442	165565467	184142759	89.9114594019958	-0.000015099769431		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSpooled	gDNA	all mappable	p<1e-3	208459		all lengths	No	No	4/15/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	gDNA upsampled	184142759	184142731	165565467	111.220494428346	0.00001520559383		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSpooled	gDNA	all mappable	p<1e-3	209954		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag:beta)	CSHL/1stpaper	NS downsampled	165565442	165565467	184142759	89.9114594019958	-0.000015099769431	15961	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSpooled	gDNA	all mappable	q<0.001	163281		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag:beta)	CSHL/1stpaper	NS downsampled	165565442	165565467	184142759	89.9114594019958	-0.000015099769431	15961	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG

Sample	Control	Read Processing (mapq, filterdup, etc)	Cut Off	Number of Peaks	Parameter Notes	peak length filtering	chrY,chrM peaks removed?	Other filtering?	Date	MACS2 Version	Purpose	If down(up) sampling, was treat or control down(up) sampled?	num treatment Reads	num Control Reads	num reads down (up) sampled from	percent of reads sampled	Percent Difference: (numTreat - numControl) /numTreat	Seed for reproducing downsampling	Command
LexoG0Rep1	gDNA	all mappable	p<1e-3	138516		all lengths	No	No	4/14/2013	macs2 2.0.10.20120913 (tag:beta)	ASBMB	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep1	gDNA	all mappable	p<1e-3	152908		all lengths	No	No	7/1/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep1	gDNA	all mappable	p<1e-3	64721		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
LexoG0Rep1	gDNA	all mappable	p<1e-5	97154		all lengths	No	No	7/1/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-5 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3
LexoG0Rep1	gDNA	all mappable	p<1e-5	38086		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
LexoG0Rep1	gDNA	all mappable	q<0.05	175130		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep1	gDNA	all mappable	q<0.01	140297		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep1	gDNA	all mappable	q<0.001	110985		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	57272428	57272406	165565467	34.5919997918406	0.000038412899135	11073	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep1	gDNA	MAPQ>=30	p<1e-3	136247		all lengths	No	No	7/7/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	41788662	41788638	144804056	28.8587482660016	0.000057431845987	32294	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep1	gDNA	MAPQ>=30	p<1e-5	90511		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	41788662	41788638	144804056	28.8587482660016	0.000057431845987	32294	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3
LexoG0Rep1	gDNA	MAPQ>=30	q<0.05	152424		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	41788662	41788638	144804056	28.8587482660016	0.000057431845987	32294	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_05/$outputPrefix05 -q $CUTOFF05 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep1	gDNA	MAPQ>=30	q<0.01	124199		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	41788662	41788638	144804056	28.8587482660016	0.000057431845987	32294	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_01/$outputPrefix01 -q $CUTOFF01 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep1	gDNA	MAPQ>=30	q<0.001	99266		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	41788662	41788638	144804056	28.8587482660016	0.000057431845987	32294	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_001/$outputPrefix001 -q $CUTOFF001 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	all mappable	p<1e-3	226868		all lengths	No	No	6/28/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep2	gDNA	all mappable	p<1e-3	112443		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
LexoG0Rep2	gDNA	all mappable	p<1e-5	177515		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	all mappable	p<1e-5	89017		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
LexoG0Rep2	gDNA	all mappable	q<0.05	252247		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	all mappable	q<0.01	225367		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	all mappable	q<0.001	194316		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	132316521	132316488	165565467	79.9179263632313	0.000024940196244	20984	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	MAPQ>=30	p<1e-5	167433		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	105892838	105892816	144804056	73.1283493882243	0.000020775720451	32387	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-5 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep2	gDNA	MAPQ>=30	q<0.05	243149		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	105892838	105892816	144804056	73.1283493882243	0.000020775720451	32387	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_05/$outputPrefix05 -q $CUTOFF05 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	MAPQ>=30	q<0.01	213194		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	105892838	105892816	144804056	73.1283493882243	0.000020775720451	32387	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_01/$outputPrefix01 -q $CUTOFF01 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep2	gDNA	MAPQ>=30	q<0.001	184824		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	105892838	105892816	144804056	73.1283493882243	0.000020775720451	32387	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_001/$outputPrefix001 -q $CUTOFF001 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	all mappable	p<1e-3	784465	--call-summits used	all lengths	No	No	6/27/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg --call-summits
LexoG0Rep3	gDNA	all mappable	p<1e-3	215301		all lengths	No	No	6/27/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep3	gDNA	all mappable	p<1e-3	103674		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	awk used to select peaks of given length from the 7/2/13 p<1e-3 set
LexoG0Rep3	gDNA	all mappable	p<1e-5	168333		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	all mappable	p<1e-5	81239		min750,max10kb	chrY/chrM removed	No	7/12/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	awk used to select peaks of given length from the 7/2/13 p<1e-5 set
LexoG0Rep3	gDNA	all mappable	q<0.05	238955		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	all mappable	q<0.01	210357		all lengths	No	No	7/9/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	all mappable	q<0.001	183916		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	115535011	115534991	165565467	69.7820584772065	0.00001731076998	14960	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	MAPQ>=30	p<1e-5	53389		all lengths	No	No	7/8/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	2820	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
LexoG0Rep3	gDNA	MAPQ>=30	q<0.05	231221		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	2820	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_05/$outputPrefix05 -q $CUTOFF05 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	MAPQ>=30	q<0.01	204451		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	2820	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_01/$outputPrefix01 -q $CUTOFF01 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
LexoG0Rep3	gDNA	MAPQ>=30	q<0.001	176258		all lengths	No	No	7/10/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled gDNA	40553346	40553321	144804056	28.0056526869662	0.000061647194291	2820	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n level_001/$outputPrefix001 -q $CUTOFF001 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
Lexo pooled	gDNA	all mappable	p<1e-3	225803		all lengths	No	No	7/29/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled Lexo pool	165565447	165565467	305123960	54.2616997367234	-0.00001207981518	448	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
Lexo pooled	gDNA	all mappable	q<0.001	197173		all lengths	No	No	7/25/2013	macs2 2.0.10.20130415 (tag:beta)	CSHL/1stpaper	downsampled Lexo pool	165565447	165565467	305123960	54.2616997367234	-0.00001207981518	448	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		

numPeaks - MCF7 NS-seq - MACS2

Sample	Control	Read Processing (mapq, filterdup, etc)	Cut Off	Number of Peaks	Parameter Notes	peak length filtering	chrY,chrM peaks removed?	Other filtering?	Date	MACS2 Version	Purpose	If down(up) sampling, was treat or control down(up) sampled?	num treatment Reads	num Control Reads	num reads down (up) sampled from	percent of reads sampled	Percent Difference: (numTreat - numControl) /numTreat	Seed for reproducing downsampling	Command
NSRep1	LexoG0 Rep1	all mappable	p<1e-3	21114		all lengths	No	No	4/15/2013	macs2 2.0.10.20120913 (tag.beta)	ASBMB	downsampled LexoRep1	45124630	45124606	57272428	78.7894063090882	0.000053186031664		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep1	Lexo pool	all mappable	p<1e-3	32756		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	45124630	45124602	305123960	14.7889408619369	0.000062050370275	793	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep1	Lexo pool	all mappable	q<0.001	4519		all lengths	No	No	7/16/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	45124630	45124602	305123960	14.7889408619369	0.000062050370275	793	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
NSRep2	LexoG0 Rep1	all mappable	p<1e-3	63159		all lengths	No	No	4/15/2013	macs2 2.0.10.20120913 (tag.beta)	ASBMB	downsampled NS Rep2	57272402	57272428	71264584	80.3658868901196	-0.000045397083223		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep2	Lexo pool	all mappable	p<1e-3	55920		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	71264584	71264558	305123960	23.3559363872965	0.000036483760292	31659	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep2	Lexo pool	all mappable	q<0.001	51654		all lengths	No	No	7/16/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	71264584	71264558	305123960	23.3559363872965	0.000036483760292	31659	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
NSRep3	LexoG0 Rep1	all mappable	p<1e-3	81640		all lengths	No	No	4/15/2013	macs2 2.0.10.20120913 (tag.beta)	ASBMB	downsampled NS Rep3	57272405	57272428	67753545	84.5305260411097	-0.000040158956133		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NSRep3	Lexo pool	all mappable	p<1e-3	85641		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	67753545	67753527	305123960	22.2052463529904	0.000026566875578	2336	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NSRep3	Lexo pool	all mappable	q<0.001	68792		all lengths	No	No	7/16/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	67753545	67753527	305123960	22.2052463529904	0.000026566875578	2336	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
																#DIV/0!	#DIV/0!		
NSpooled	LexoG0 Rep1	all mappable	p<1e-3	87812		all lengths	No	No	4/18/2013	macs2 2.0.10.20120913 (tag.beta)	ASBMB	upsampled Lexo Rep1	184142759	184142729	57272428	321.520730708326	0.000016291707674		macs2 callpeak -t $Treatment -c $Control -f BED -g hs -n $outputPrefix -p 1e-3 -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 --bdg
NS pool	Lexo pool	all mappable	p<1e-3	88362		all lengths	No	No	7/17/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	184142759	184142742	305123960	60.3501416276847	0.000009231967682	20042	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -p $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG
NS pool	Lexo pool	all mappable	q<0.001	67867		all lengths	No	No	7/16/2013	macs2 2.0.10.20130712 (tag.beta)	CSHL/1st paper	downsampled lexo pool	184142759	184142742	305123960	60.3501416276847	0.000009231967682	20042	macs2 callpeak -t $Treatment -c $Control -f BED --down-sample --seed $SEED -g hs -n $outputPrefix -q $CUTOFF -s 50 --keep-dup all --nomodel --shiftsize=175 --slocal 5000 --llocal 50000 --verbose 3 $BDG

	Lexo Pooled	Lexo Rep1	Lexo Rep2	Lexo Rep3	NS-gDNA Pool	NS-gDNA Rep1	NS-gDNA Rep2	NS-gDNA Rep3	NS-lexo Pool	NS-lexo Rep1	NS-lexo Rep2	NS-lexo Rep3
Lexo Pooled		0.84553	0.975653	0.971312	0.557402	0.714478	0.569375	0.325473	0.171399	0.375128	0.133378	0.0858222
Lexo Rep1			0.789778	0.781041	0.682616	0.737715	0.65975	0.506009	0.327537	0.439367	0.272815	0.252542
Lexo Rep2				0.950533	0.511703	0.691209	0.527984	0.275681	0.133805	0.361939	0.0979856	0.0458444
Lexo Rep3					0.507415	0.672902	0.528688	0.27693	0.13193	0.332397	0.101286	0.0514567
NS-gDNA Pooled						0.783419	0.927668	0.924786	0.781847	0.62966	0.701933	0.724164
NS-gDNA Rep1							0.675875	0.571822	0.422503	0.741351	0.296377	0.292583
NS-gDNA Rep2								0.806056	0.703584	0.481799	0.736385	0.618532
NS-gDNA Rep3									0.849206	0.523613	0.732086	0.865571
NS-lexo Pooled										0.652201	0.921222	0.952993
NS-lexo Rep1											0.488097	0.494197
NS-lexo Rep2												0.845581
NS-lexo Rep3												

	Lexo Pooled	Lexo Rep1	Lexo Rep2	Lexo Rep3	NS-gDNA Pool	NS-gDNA Rep1	NS-gDNA Rep2	NS-gDNA Rep3	NS-lexo Pool	NS-lexo Rep1	NS-lexo Rep2	NS-lexo Rep3
Lexo Pooled		0.724043	0.969306	0.968379	0.357286	0.532964	0.435164	0.209907	0.0391832	0.123031	0.0220685	0.00898538
Lexo Rep1			0.61876	0.620407	0.672747	0.759268	0.671822	0.527576	0.223922	0.344102	0.177025	0.150357
Lexo Rep2				0.949358	0.271078	0.451778	0.3617	0.13656	0.00657257	0.0794766	-0.0047056	-0.0123113
Lexo Rep3					0.276412	0.455444	0.361749	0.142022	0.013387	0.0812468	0.0020833	-0.00744964
NS-gDNA Pooled						0.854673	0.952862	0.931217	0.612074	0.544475	0.537175	0.516007
NS-gDNA Rep1							0.802021	0.67504	0.305339	0.621268	0.234927	0.19006
NS-gDNA Rep2								0.853055	0.568699	0.46612	0.536337	0.46428
NS-gDNA Rep3									0.785144	0.490705	0.689196	0.747458
NS-lexo Pooled										0.442739	0.941841	0.953558
NS-lexo Rep1											0.363807	0.315428
NS-lexo Rep2												0.879957
NS-lexo Rep3												

MCF7datasetSignalCorrelations_NS_LexoG0 - -log10(qVal)

	Lexo Pooled	Lexo Rep1	Lexo Rep2	Lexo Rep3	NS-gDNA Pool	NS-gDNA Rep1	NS-gDNA Rep2	NS-gDNA Rep3	NS-lexo Pool	NS-lexo Rep1	NS-lexo Rep2	NS-lexo Rep3
Lexo Pooled		0.693613	0.968296	0.967111	0.349159	0.497608	0.409816	0.207041	0.0447436	0.143733	0.0371851	0.0138325
Lexo Rep1			0.578881	0.584371	0.687005	0.764301	0.680105	0.540661	0.220356	0.401475	0.188621	0.136825
Lexo Rep2				0.947634	0.259957	0.407898	0.330944	0.130549	0.0119374	0.0867982	0.0084334	-0.0066134
Lexo Rep3					0.265795	0.415464	0.332956	0.136433	0.01848	0.0943613	0.0152868	-0.0022265
NS-gDNA Pooled						0.870376	0.956168	0.935102	0.605608	0.585858	0.539715	0.503397
NS-gDNA Rep1							0.821299	0.693376	0.301037	0.638573	0.250944	0.1763
NS-gDNA Rep2								0.854943	0.546617	0.525043	0.522164	0.431503
NS-gDNA Rep3									0.772421	0.492741	0.672888	0.728553
NS-lexo Pooled										0.387038	0.935265	0.950722
NS-lexo Rep1											0.332599	0.229176
NS-lexo Rep2												0.855142
NS-lexo Rep3												

Sample1	Sample2	FE	pVal	qVal	Subtract	Pileup	ControlLambda
LexoRep1-gDNA	LexoRep2-gDNA	0.789778	0.61876	0.578881	0.785103	0.848402	0.969207
LexoRep1-gDNA	LexoRep3-gDNA	0.781041	0.620407	0.584371	0.786015	0.845567	0.966883
LexoRep1-gDNA	Lexopool-gDNA	0.84553	0.724043	0.693613	0.846994	0.89562	0.972348
LexoRep1-gDNA	NSpool-gDNA	0.682616	0.672747	0.687005	0.694771	0.792333	0.972348
LexoRep1-gDNA	NSpool-lexo	0.327537	0.223922	0.220356	-0.113027	0.79327	0.55914
LexoRep1-gDNA	NSRep1-gDNA	0.737715	0.759268	0.764301	0.768571	0.846324	0.93847
LexoRep1-gDNA	NSRep1-lexo	0.439367	0.344102	0.401475	0.11193	0.846324	0.552331
LexoRep1-gDNA	NSRep2-gDNA	0.65975	0.671822	0.680105	0.667474	0.786284	0.955596
LexoRep1-gDNA	NSRep2-lexo	0.272815	0.177025	0.188621	-0.199629	0.786284	0.555607
LexoRep1-gDNA	NSRep3-gDNA	0.506009	0.527576	0.540661	0.491742	0.607408	0.954058
LexoRep1-gDNA	NSRep3-lexo	0.252542	0.150357	0.136825	-0.140048	0.607408	0.555296
LexoRep2-gDNA	LexoRep3-gDNA	0.950533	0.949358	0.947634	0.961696	0.972185	0.989656
LexoRep2-gDNA	Lexopool-gDNA	0.975653	0.969306	0.968296	0.979278	0.984891	0.996042
LexoRep2-gDNA	NSpool-gDNA	0.511703	0.271078	0.259957	0.530516	0.646246	0.996042
LexoRep2-gDNA	NSpool-lexo	0.133805	0.00657257	0.0119374	-0.386851	0.647053	0.568844
LexoRep2-gDNA	NSRep1-gDNA	0.691209	0.451778	0.407898	0.69102	0.776521	0.95846
LexoRep2-gDNA	NSRep1-lexo	0.361939	0.0794766	0.0867982	-0.10894	0.77652	0.561794
LexoRep2-gDNA	NSRep2-gDNA	0.527984	0.3617	0.330944	0.551925	0.675066	0.977041
LexoRep2-gDNA	NSRep2-lexo	0.0979856	-0.00470561	0.0084334	-0.431337	0.675066	0.565172
LexoRep2-gDNA	NSRep3-gDNA	0.275681	0.13656	0.130549	0.27168	0.406054	0.975358
LexoRep2-gDNA	NSRep3-lexo	0.0458444	-0.0123113	-0.0066134	-0.414567	0.406055	0.564833
LexoRep3-gDNA	Lexopool-gDNA	0.971312	0.968379	0.967111	0.977614	0.983334	0.993282
LexoRep3-gDNA	NSpool-gDNA	0.507415	0.276412	0.265795	0.534361	0.64603	0.993282
LexoRep3-gDNA	NSpool-lexo	0.13193	0.013387	0.01848	-0.388915	0.646839	0.567924
LexoRep3-gDNA	NSRep1-gDNA	0.672902	0.455444	0.415464	0.687562	0.770697	0.956232
LexoRep3-gDNA	NSRep1-lexo	0.332397	0.0812468	0.0943613	-0.119387	0.770696	0.56089
LexoRep3-gDNA	NSRep2-gDNA	0.528688	0.361749	0.332956	0.558227	0.675696	0.974632
LexoRep3-gDNA	NSRep2-lexo	0.101286	0.0020833	0.0152868	-0.431438	0.675696	0.564256
LexoRep3-gDNA	NSRep3-gDNA	0.27693	0.142022	0.136433	0.278128	0.409026	0.972984
LexoRep3-gDNA	NSRep3-lexo	0.0514567	-0.00744964	-0.00222654	-0.413092	0.409027	0.56393
Lexopool-gDNA	NSpool-gDNA	0.557402	0.357286	0.349159	0.57951	0.688639	1
Lexopool-gDNA	NSpool-lexo	0.171399	0.0391832	0.0447436	-0.351774	0.689491	0.570331
Lexopool-gDNA	NSRep1-gDNA	0.714478	0.532964	0.497608	0.7262	0.805749	0.961461
Lexopool-gDNA	NSRep1-lexo	0.375128	0.123031	0.143733	-0.0777374	0.805749	0.563227
Lexopool-gDNA	NSRep2-gDNA	0.569375	0.435164	0.409816	0.594041	0.712062	0.980335
Lexopool-gDNA	NSRep2-lexo	0.133378	0.0220685	0.0371851	-0.404041	0.712062	0.566622
Lexopool-gDNA	NSRep3-gDNA	0.325473	0.209907	0.207041	0.322455	0.45413	0.978613
Lexopool-gDNA	NSRep3-lexo	0.0858222	0.00898538	0.0138325	-0.379053	0.454131	0.566286
NSpool-gDNA	NSpool-lexo	0.781847	0.612074	0.605608	0.460101	0.99878	0.570331
NSpool-gDNA	NSRep1-gDNA	0.783419	0.854673	0.870376	0.820555	0.879337	0.961461
NSpool-gDNA	NSRep1-lexo	0.62966	0.544475	0.585858	0.482423	0.879337	0.563227

NSpool-gDNA	NSRep2-gDNA	0.927668	0.952862	0.956168	0.929524	0.954129	0.980335
NSpool-gDNA	NSRep2-lexo	0.701933	0.537175	0.539715	0.348914	0.954129	0.566622
NSpool-gDNA	NSRep3-gDNA	0.924786	0.931217	0.935102	0.913091	0.925364	0.978613
NSpool-gDNA	NSRep3-lexo	0.724164	0.516007	0.503397	0.43046	0.925364	0.566286
NSpool-lexo	NSRep1-gDNA	0.422503	0.305339	0.301037	0.12849	0.880411	0.554042
NSpool-lexo	NSRep1-lexo	0.652201	0.442739	0.387038	0.752172	0.880411	0.981605
NSpool-lexo	NSRep2-gDNA	0.703584	0.568699	0.546617	0.359685	0.955308	0.562342
NSpool-lexo	NSRep2-lexo	0.921222	0.941841	0.935265	0.923344	0.955308	0.988494
NSpool-lexo	NSRep3-gDNA	0.849206	0.785144	0.772421	0.639898	0.926483	0.561672
NSpool-lexo	NSRep3-lexo	0.952993	0.953558	0.950722	0.952472	0.926483	0.987859
NSRep1-gDNA	NSRep1-lexo	0.741351	0.621268	0.638573	0.505222	1	0.547349
NSRep1-gDNA	NSRep2-gDNA	0.675875	0.802021	0.821299	0.720047	0.813161	0.945364
NSRep1-gDNA	NSRep2-lexo	0.296377	0.234927	0.250944	-0.0238999	0.813161	0.550587
NSRep1-gDNA	NSRep3-gDNA	0.571822	0.67504	0.693376	0.595377	0.684448	0.94388
NSRep1-gDNA	NSRep3-lexo	0.292583	0.19006	0.1763	0.0355339	0.684448	0.550259
NSRep1-lexo	NSRep2-gDNA	0.481799	0.46612	0.525043	0.317887	0.813161	0.555458
NSRep1-lexo	NSRep2-lexo	0.488097	0.363807	0.332599	0.602373	0.813162	0.974487
NSRep1-lexo	NSRep3-gDNA	0.523613	0.490705	0.492741	0.480752	0.684449	0.554837
NSRep1-lexo	NSRep3-lexo	0.494197	0.315428	0.229176	0.622247	0.68445	0.973871
NSRep2-gDNA	NSRep2-lexo	0.736385	0.536337	0.522164	0.362584	1	0.55881
NSRep2-gDNA	NSRep3-gDNA	0.806056	0.853055	0.854943	0.788873	0.827478	0.961402
NSRep2-gDNA	NSRep3-lexo	0.618532	0.46428	0.431503	0.296663	0.827478	0.558445
NSRep2-lexo	NSRep3-gDNA	0.732086	0.689196	0.672888	0.491142	0.827479	0.558108
NSRep2-lexo	NSRep3-lexo	0.845581	0.879957	0.855142	0.850364	0.827479	0.980641
NSRep3-gDNA	NSRep3-lexo	0.865571	0.747458	0.728553	0.68541	1	0.557798

Lexo G0 Replicates and pooled sample

Sample1	Sample2	FE	pVal	qVal	Subtract	Pileup	ControlLambda
LexoRep1-gDNA	LexoRep2-gDNA	0.789778	0.61876	0.578881	0.785103	0.848402	0.969207
LexoRep1-gDNA	LexoRep3-gDNA	0.781041	0.620407	0.584371	0.786015	0.845567	0.966883
LexoRep2-gDNA	LexoRep3-gDNA	0.950533	0.949358	0.947634	0.961696	0.972185	0.989656
LexoRep1-gDNA	Lexopool-gDNA	0.84553	0.724043	0.693613	0.846994	0.89562	0.972348
LexoRep2-gDNA	Lexopool-gDNA	0.975653	0.969306	0.968296	0.979278	0.984891	0.996042
LexoRep3-gDNA	Lexopool-gDNA	0.971312	0.968379	0.967111	0.977614	0.983334	0.993282

gDNA-controlled NS Replicates and pooled sample

Sample1	Sample2	FE	pVal	qVal	Subtract	Pileup	ControlLambda
NSRep1-gDNA	NSRep2-gDNA	0.675875	0.802021	0.821299	0.720047	0.813161	0.945364
NSRep1-gDNA	NSRep3-gDNA	0.571822	0.67504	0.693376	0.595377	0.684448	0.94388
NSRep2-gDNA	NSRep3-gDNA	0.806056	0.853055	0.854943	0.788873	0.827478	0.961402
NSpool-gDNA	NSRep1-gDNA	0.783419	0.854673	0.870376	0.820555	0.879337	0.961461
NSpool-gDNA	NSRep2-gDNA	0.927668	0.952862	0.956168	0.929524	0.954129	0.980335
NSpool-gDNA	NSRep3-gDNA	0.924786	0.931217	0.935102	0.913091	0.925364	0.978613

Lexo-controlled NS Replicates and pooled sample

Sample1	Sample2	FE	pVal	qVal	Subtract	Pileup	ControlLambda
NSRep1-lexo	NSRep2-lexo	0.488097	0.363807	0.332599	0.602373	0.813162	0.974487
NSRep1-lexo	NSRep3-lexo	0.494197	0.315428	0.229176	0.622247	0.68445	0.973871
NSRep2-lexo	NSRep3-lexo	0.845581	0.879957	0.855142	0.850364	0.827479	0.980641
NSpool-lexo	NSRep1-lexo	0.652201	0.442739	0.387038	0.752172	0.880411	0.981605
NSpool-lexo	NSRep2-lexo	0.921222	0.941841	0.935265	0.923344	0.955308	0.988494
NSpool-lexo	NSRep3-lexo	0.952993	0.953558	0.950722	0.952472	0.926483	0.987859

Window=100kb, stepSize=100kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.78569014157145	0.984769059730476	0.978965969517742	0.631439957573922	0.010683276282506	0.345595948705346	-0.225942951983836	-0.112750673754164	-0.258898580423667	0.504978515103589	-0.265725616222639
LexoRep1			0.797594573205734	0.823619979340527	0.836426662298997	0.030229199138417	0.602012370463512	-0.1879681023934	-0.054755046032047	-0.212070612062567	0.690722414894663	-0.218437924188718
LexoRep2				0.977934538698587	0.640994568367816	0.003594804050583	0.357901453743449	-0.22384344443563	-0.115576516465621	-0.259147009430784	0.512097798670524	-0.266864264730152
LexoRep3					0.663667084236949	0.012765468404613	0.383573703487132	-0.228859935893412	-0.1211312681431	-0.264639902914906	0.531309081271953	-0.271065929380908
NSRep1-gDNA						0.101895963430887	0.687132193326882	-0.156491941534954	0.051251305896398	-0.178341887293044	0.86113566346463	-0.156223533654623
NSRep1-lexo							0.154288252315288	0.211405432985226	0.25189651211163	0.192598134179896	0.143845002304524	0.244961412202877
NSRep2-gDNA								0.410289489644829	0.416059176320858	0.30458680670668	0.769139017288429	0.301194794876909
NSRep2-lexo									0.622778621375684	0.730417614102143	0.061898867820156	0.728178278590855
NSRep3-gDNA										0.811927603574207	0.377533593847262	0.783296022683351
NSRep3-lexo											0.09847507586529 7	0.926307423920325
NSpool-gDNA												0.124729747915347
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.778572132965817	0.968699361054858	0.963760442454606	0.657647191409398	0.052735185657039	0.231287262008367	-0.243283660398255	-0.068668037413384	-0.253616808562338	0.386253885284427	-0.249750693551336
LexoRep1			0.772160857057762	0.782626532901487	0.733634470470805	0.121391965250108	0.324001449723028	-0.193689629373216	-0.027553504695084	-0.206723887676371	0.440238095149952	-0.194423934140592
LexoRep2				0.952010242449653	0.655331850101909	0.051811275231525	0.23877224421465	-0.234122539566043	-0.062690683290488	-0.245956050981159	0.390511303800798	-0.243439347506344
LexoRep3					0.659802336725253	0.051024991290046	0.228563765938542	-0.250835679445539	-0.082675238377865	-0.26428368839924	0.377323048086124	-0.261112533308248
NSRep1-gDNA						0.233864415509387	0.397960623099061	-0.135642623794551	0.114468743623273	-0.128758090258718	0.624031440127568	-0.073137841451917
NSRep1-lexo							0.147664422812574	0.131343565420984	0.168804520120369	0.104805000746605	0.182365601523283	0.169626146077534
NSRep2-gDNA								0.58987614048633	0.565734889442824	0.49200215946202	0.642791386106908	0.487101890072763
NSRep2-lexo									0.584909455937884	0.740792105791 92	0.198494543187682	0.73167500617056
NSRep3-gDNA										0.801059901695332	0.596795236742397	0.787348114680829
NSRep3-lexo											0.316206614177084	0.909093240099177
NSpool-gDNA												0.373308283029781
NSpool-lexo												

Window=100kb, stepSize=25kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.785840438535835	0.98474104101105	0.978827630224493	0.632447005885261	0.008905054697214	0.344509881457207	-0.224804444261724	-0.113284155168903	-0.258533398658315	0.50412533678337	-0.264965198184888
LexoRep1			0.797130757472103	0.823497039966692	0.836060402009599	0.030239738601448	0.60061031741153	-0.186261363498568	-0.053327324157252	-0.210777519221017	0.689172582603028	-0.217455869379024
LexoRep2				0.977810134023734	0.641535968554449	0.002170478037208	0.356668449248923	-0.222196259439866	-0.115729047199555	-0.258737182257737	0.510901822501126	-0.266448820916488
LexoRep3					0.664608003818104	0.010604951642226	0.382508997779389	-0.227922244096402	-0.121640440967707	-0.264277788780929	0.530452213906512	-0.270511859316298
NSRep1-gDNA						0.100867142964591	0.686052962616596	-0.156580895865691	0.050950835929498	-0.178208937194106	0.860721398514787	-0.156203339200128
NSRep1-lexo							0.153841003295457	0.20680388311537	0.247729900952247	0.193264316637366	0.141686867777483	0.244638788832798
NSRep2-gDNA								0.411948387728018	0.41611154669148	0.306742199376642	0.768446324408572	0.303641129594071
NSRep2-lexo									0.622764182913162	0.731888747245107	0.062601807032377	0.730944370687475
NSRep3-gDNA										0.814891360985829	0.378745568598176	0.785878985251648
NSRep3-lexo											0.100272627060749	0.925855144411856
NSpool-gDNA												0.126286144281261
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.778756015668262	0.968831919483506	0.96355815178524	0.657506019540772	0.050576863782078	0.231269115123495	-0.241629758993748	-0.066544848420874	-0.251137723086673	0.388379753697222	-0.247103662343394
LexoRep1			0.772438010246604	0.782786361611472	0.735277942918469	0.12405114398195	0.325444567452702	-0.192162128512799	-0.026245527760785	-0.205667274420134	0.442778096735999	-0.192487583197104
LexoRep2				0.952073161699451	0.656175885013487	0.049553736468819	0.239001843158288	-0.231869732623293	-0.059807012313426	-0.243019289379142	0.393017252513717	-0.241029640495794
LexoRep3					0.660594229707126	0.050058979330592	0.227582907627141	-0.250660939735231	-0.080946879618625	-0.26228946271626	0.379568339778605	-0.25875579489109
NSRep1-gDNA						0.232875454836632	0.398239124656846	-0.13477385014853	0.114871880968526	-0.129155583256224	0.626573578868 86	-0.072946045434167
NSRep1-lexo							0.148646669450977	0.12826203742543	0.164921825150069	0.102985393718297	0.179538279692752	0.168710503006064
NSRep2-gDNA								0.590960873024713	0.567572310301042	0.492815219584534	0.644236012019838	0.489279435832137
NSRep2-lexo									0.585439408661765	0.742105615645504	0.198829799540724	0.734023797931553
NSRep3-gDNA										0.801252769656361	0.59610252689745	0.786985587098577
NSRep3-lexo											0.315068565778546	0.909127573345145
NSpool-gDNA												0.371737278011408
NSpool-lexo												

RATIOS

This shows us that it does not make a difference if we use overlapping windows or just partition the genome. The correlation scores are essentially the same for this particular analysis. Shown is just correlation(100kb bins, 100kb steps)/correlation(100kb bins, 25kb steps)

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.99980874366218	1.00002845288077	1.00014133161854	0.998407695345273	1.19968676731973	1.00315249955544	1.00506443600726	0.995290767592841	1.00141251291805	1.00169239325614	1.00286987892357
LexoRep1			1.0005818565264	1.00014928939373	1.0004380787423	0.999651469770625	1.0023343805648	1.00916313970205	1.0267728016989	1.00613487076956	1.00224883045373	1.00451611084354
LexoRep2				1.00012722784365	0.999156087556785	1.65622687212591	1.00345700466953	1.00741319858361	0.998682001298507	1.00158395159703	1.00234091192618	1.00155918803557
LexoRep3					0.99858424879666	1.20372716776797	1.00278347885651	1.00411408636628	0.995814115597113	1.0013702026782	1.00161535260477	1.00204822836977
NSRep1-gDNA						1.01019975817752	1.0015731011585	0.999431895377497	1.00589725293842	1.00074603496902	1.00048129969878	1.00012928311647
NSRep1-lexo							1.00290721595836	1.02225079046165	1.01681916935891	0.996552998147506	1.01523171879578	1.00131877439231

NSRep2-gDNA								0.995973043874894	0.999874143433321	0.992973276339734	1.00090141999233	0.991943335474899
NSRep2-lexo									1.00002318447804	0.997989949772419	0.988771263234343	0.996215728299517
NSRep3-gDNA										0.996363003028972	0.996800029224369	0.9967132820488
NSRep3-lexo											0.982073360914709	1.000488499212
NSpool-gDNA												0.987675636351304
NSpool-lexo												
Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.999763876363398	0.999863177063036	1.00020994131905	1.00021470810066	1.04267409470584	1.00007846652962	1.00684477529339	1.03190613613065	1.00987141814139	0.994526315049748	1.01071222976964
LexoRep1			0.999641196853126	0.99979582077841	0.997764828302695	0.978563851597943	0.995565703428484	1.00794902134067	1.04983618337648	1.00513748849551	0.994263488630601	1.01005961481425
LexoRep2				0.99993391342984	0.998713706293	1.04555738726436	0.99903934237241	1.00971583016577	1.04821626872029	1.0120844794235	0.993623820081966	1.0099975546808
LexoRep3					0.998801241448591	1.01929747614458	1.00430989445397	1.00069711583501	1.02135176510056	1.0076031483016	0.9940846180854	1.00910796381642
NSRep1-gDNA						1.00424673640873	0.999300667512201	1.0064461588436	0.996490548062293	0.996922370775738	0.99594279295038	1.00262928602378
NSRep1-lexo							0.993392070996068	1.02402525374935	1.023542638864	1.01766859321124	1.01574773822813	1.00542730331044
NSRep2-gDNA								0.998164459631936	0.996762666492234	0.998350172457743	0.997757613846514	0.995549484405224
NSRep2-lexo									0.999094777843718	0.998482149922377	0.998313852582378	0.996800114972278
NSRep3-gDNA										0.99975929198833	1.0011620649363	1.00046065339467
NSRep3-lexo											1.0036120658236	0.999962234952525
NSpool-gDNA												1.00422611643034
NSpool-lexo												
Added 7/30/13												
Window=10kb, stepSize=5kb												
Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.657264342100861	0.934920317302157	0.921326147246424	0.514463084550608	0.036040180862913	0.253870590664468	-0.0578628728405	0.014303133491125	-0.042076650176422	0.370060838436823	-0.05030833581226
LexoRep1			0.665831350604074	0.696844075692386	0.68898060479129	0.062748355823912	0.438363990931231	-0.026110808923071	0.0287236901349	-0.017900520076101	0.509795467945734	-0.022911010049103
LexoRep2				0.90997447540568	0.522502716432692	0.030811751618476	0.261609954717295	-0.05572589050938	0.012753292047942	-0.043462156060687	0.376171214659096	-0.052651785980729
LexoRep3					0.543217036789468	0.037668791781788	0.277171143688251	-0.062609309719905	-0.000400947354009	-0.052839504297948	0.386613742545584	-0.060204179555708
NSRep1-gDNA						0.134413136320575	0.496498081683875	0.009412113739235	0.121627908193285	0.018563834779722	0.674893937702785	0.036063704795119
NSRep1-lexo							0.138753998727597	0.157840102387629	0.147330460068315	0.13409354197859	0.114424137979188	0.176949653603819
NSRep2-gDNA								0.607555125478609	0.458370923566244	0.463713864996224	0.645966480663666	0.474368898276692
NSRep2-lexo									0.516447301296178	0.596642666166795	0.273742999806116	0.620964972363287
NSRep3-gDNA										0.805616644619953	0.547234708290176	0.756506645414226
NSRep3-lexo											0.401746076708265	0.885190489896976
NSpool-gDNA												0.417963839799195
NSpool-lexo												
Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.683326093099801	0.931227593888123	0.920199973989484	0.544826573769273	0.063056125860336	0.258922792377552	-0.042878342915639	0.063649394186563	-0.023246885149449	0.335334715368675	-0.034241511814184
LexoRep1			0.676129565297424	0.692714008619526	0.61282074789125	0.102326705073495	0.362842565798418	-0.011845530765598	0.080321164789614	-0.000414594298619	0.395555035054136	-0.007013753287551
LexoRep2				0.897535036911201	0.542640266330175	0.060412440424757	0.262126470624537	-0.038369154563706	0.066200655445165	-0.018584204603326	0.337392546008041	-0.030780093179648
LexoRep3					0.547774971778668	0.062822676216682	0.2590762505934	-0.049832602637254	0.045943370117364	-0.036681557120861	0.325647983181013	-0.04703261974368
NSRep1-gDNA						0.213477620590717	0.411362064044949	0.047174860089758	0.200022228955309	0.070090692863903	0.574785393774297	0.096731600060798
NSRep1-lexo							0.14079329676528	0.135544428654345	0.147196836605151	0.121418880662484	0.146960293056765	0.164490643423097
NSRep2-gDNA								0.67583935073274	0.607594905041622	0.606599995428765	0.676848037713271	0.602520035164435
NSRep2-lexo									0.615549274125136	0.720265411586874	0.435546590767253	0.727205249356598
NSRep3-gDNA										0.834905197701907	0.681512195521532	0.798874672160491
NSRep3-lexo											0.560126908221097	0.895824856060634
NSpool-gDNA												0.581816854001613
NSpool-lexo												

Window=100kb, stepSize=100kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.826193100569173	0.984700612338047	0.979316518596418	0.679155955305625	-0.035557179379401	0.400549437377856	-0.245108167183775	0.007115579264663	-0.290417769288983	0.538511352224799	-0.29007657224627
LexoRep1			0.828526829758066	0.850736539450549	0.834509849870514	-0.019135233822511	0.648743105654954	-0.195219806538083	0.08113602542229	-0.242577694168952	0.700017246279531	-0.243162627631355
LexoRep2				0.977387547780674	0.678041913764952	-0.046017757931021	0.404569457068038	-0.244173652931727	-0.000871281852121	-0.290481459835257	0.53433837798108	-0.292340406541176
LexoRep3					0.692810802474044	-0.050373266316384	0.428629175501442	-0.245642037473078	-0.010991502161736	-0.295823456336256	0.54471916477406	-0.298107256176361
NSRep1-gDNA						0.226347355610309	0.703399865696791	-0.213556958670887	0.257439223008302	-0.222284585959675	0.896061375601946	-0.142178465136031
NSRep1-lexo							0.09897874856822	0.109398994226561	0.385753027679238	0.164779199371382	0.298779914870165	0.36066436305693
NSRep2-gDNA								0.314638745075141	0.471048763673984	0.205744727654817	0.779138145280407	0.194301756152904
NSRep2-lexo									0.498611381853081	0.725766109498745	-0.007202097090799	0.700275338707792
NSRep3-gDNA										0.685173518244062	0.536819011384577	0.692270143153715
NSRep3-lexo											0.031373356580379	0.878742554338562
NSpool-gDNA												0.116660052306689
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.811315587180425	0.971139685804748	0.966972376799104	0.696679525045461	0.046450205945257	0.273657233069206	-0.270676323424001	0.028220016803876	-0.273539180387321	0.447875607604769	-0.247030816081056
LexoRep1			0.80251867313295	0.807246202027344	0.752272850933135	0.13409432620811	0.391396934768939	-0.186451348621995	0.099294358482728	-0.201441870024096	0.518370283354597	-0.163531701371858
LexoRep2				0.957238351616256	0.689540192116505	0.038564422432892	0.274794505891864	-0.26543433142818	0.027860634901257	-0.269462316200899	0.444448720105678	-0.245820771160139
LexoRep3					0.688711976013502	0.035405327240183	0.27259799910983	-0.27303103084282	0.011261525362195	-0.281108207653134	0.433182977788931	-0.258966045601036
NSRep1-gDNA						0.390482933084411	0.372048676898435	-0.236552197922141	0.215695039972946	-0.174067726880123	0.694566138612151	-0.054350378494718
NSRep1-lexo							0.109825060739741	0.072500292982972	0.289129899540165	0.151921606705125	0.312603159951872	0.308557009283806
NSRep2-gDNA								0.521025774761871	0.564856805429714	0.407161997482448	0.642997732093735	0.399643707374428
NSRep2-lexo									0.481500637774316	0.733963681703369	0.107759733328962	0.686196310440965
NSRep3-gDNA										0.717303104152863	0.647315036057006	0.721587684578883
NSRep3-lexo											0.243763492772086	0.872835729733461
NSpool-gDNA												0.352089900631449
NSpool-lexo												

Window=100kb, stepSize=25kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.826481927165411	0.984593198467069	0.979573564491808	0.680095796994455	-0.036312115917594	0.399780285008362	-0.243036075209391	0.007131503940414	-0.289363842683214	0.538580601308859	-0.288888931659765
LexoRep1			0.828544060346145	0.85097439467043	0.8344949692435	-0.019204732910647	0.647872211704678	-0.19231099958363	0.08235423139862	-0.239811655796459	0.699830774709744	-0.241117691673486
LexoRep2				0.97738466886064	0.678330628737472	-0.046233054960642	0.402839910894102	-0.242164800393861	-0.000946641788507	-0.289611269560709	0.533625812958023	-0.291256831338405
LexoRep3					0.693488799758232	-0.051353982411372	0.427602739901587	-0.243295605638287	-0.01087352550436	-0.294599286911794	0.544506758923095	-0.29708245190446
NSRep1-gDNA						0.226094196187092	0.702918194953897	-0.212236175066644	0.257081253589579	-0.220796232802842	0.89629039095567	-0.141405633631801
NSRep1-lexo							0.097873174335752	0.110261417751827	0.38558970342394	0.167666571566165	0.298791941233884	0.362724746424849
NSRep2-gDNA								0.318089585469435	0.471736756244103	0.208065066443561	0.779298433380372	0.196901267487273
NSRep2-lexo									0.498820622182709	0.725165652688181	-0.005578461783247	0.699589798620721
NSRep3-gDNA										0.687679882221739	0.536974289495102	0.69335428814961
NSRep3-lexo											0.033139466186471	0.878756563353882
NSpool-gDNA												0.117393154036418
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.811773318837822	0.971436708078547	0.966986355718805	0.697590487703582	0.047541936421487	0.274195955940611	-0.268544212829437	0.030036375237624	-0.271599215498947	0.449364441153045	-0.244942459571763
LexoRep1			0.802849147159924	0.807835437303407	0.752131452471433	0.135375471820628	0.392399927644351	-0.18391214556329	0.102248804875532	-0.200840019436334	0.519914517198324	-0.162224327356674
LexoRep2				0.957742216347206	0.690369138314761	0.039702687502368	0.275743450181469	-0.263221810175749	0.030494867733728	-0.266814032439356	0.446279561277548	-0.243125608920217
LexoRep3					0.690291794407522	0.03685187521899	0.272921115198822	-0.271210823419948	0.0131277097926	-0.279624620870503	0.435245872547307	-0.256847061217139
NSRep1-gDNA						0.393765365964468	0.373782694313105	-0.233506296835618	0.217180289032671	-0.173851646147876	0.696004812988588	-0.052999619761532
NSRep1-lexo							0.112912619546711	0.073730747303218	0.289685140677083	0.148980675877114	0.316093383919491	0.307453107184001
NSRep2-gDNA								0.522528257033864	0.564914071245629	0.408327235467323	0.643472293337263	0.401122916334695
NSRep2-lexo									0.480763606635751	0.734396141625791	0.108085210995771	0.686124592628354
NSRep3-gDNA										0.715449058919561	0.647212318034678	0.720562984099118
NSRep3-lexo											0.241784460244438	0.871796548654059
NSpool-gDNA												0.350751005064152
NSpool-lexo												

RATIOS

This shows us that it does not make a difference if we use overlapping windows or just partition the genome. The correlation scores are essentially the same for this particular analysis. Shown is just correlation(100kb bins, 100kb steps)/correlation(100kb bins, 25kb steps)

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.999650534891636	1.00010909467092	0.999737594087154	0.998618074552168	0.979209789374262	1.00192393771863	1.00852586173719	0.997766996150594	1.00364221941482	0.999871422988701	1.00411106296002
LexoRep1			0.999979203775751	0.999720490744057	1.00001783189541	0.996381147894149	1.00134423723466	1.01512553603668	0.985207730609087	1.01153421155993	1.00026645237181	1.00848106973684
LexoRep2				1.00000294553427	0.999574374264866	0.995343222942886	1.00429338336933	1.00829539443635	0.920392341326011	1.00300468374683	1.00133532712578	1.00372034261923
LexoRep3					0.999022338523096	0.980902822937237	1.00240044205538	1.00964436586775	1.01084989935681	1.00415537130892	1.000390088548	1.00344956178102
NSRep1-gDNA						1.00111970774786	1.00068524438029	1.00622317851247	1.00139243688027	1.00674084488643	0.999744485318559	1.00546535158735
NSRep1-lexo							1.01129598830294	0.9921783744228	1.00042357006385	0.982779082509935	0.99995975003988	0.994319705539181

NSRep2-gDNA								0.989151356875764	0.998541575230227	0.988848013612254	0.999794317435915	0.986797894358212
NSRep2-lexo									0.999580529913314	1.00082802709745	1.291054303254	1.00097991721495
NSRep3-gDNA										0.996355333284464	0.999710827662399	0.998436376590692
NSRep3-lexo											0.946706757551424	0.99998405813862
NSpool-gDNA												0.993755157736868
NSpool-lexo												
Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.999436133651138	0.99969424433797	0.999985543829426	0.998694129759252	0.977036474354953	0.998035263249755	1.00793951421295	0.939528041603609	1.0071427484973	0.996686801598151	1.00852590650451
LexoRep1			0.999588373447063	0.999270599866193	1.00018799700669	0.990536353481999	0.997443952445575	1.01380660886168	0.971105321021596	1.00299666664767	0.997029831265246	1.00805905030698
LexoRep2				0.999473903601251	0.998799271067824	0.971330276586236	0.996558597170738	1.00840553923307	0.913617174684202	1.00992557901596	0.99589754644683	1.01108547245143
LexoRep3					0.997711375961848	0.960746964158242	0.998816082483187	1.0067114114397	0.857843869198194	1.00530563717176	0.995260392140418	1.00824998492821
NSRep1-gDNA						0.991663987836977	0.995360894334992	1.01304419250273	0.99316121611984	1.00124290299824	0.997932953408383	1.02548619667959
NSRep1-lexo							0.97265532569021	0.983311517036634	0.998083294380857	1.01974035095959	0.988958250488065	1.00359047306406
NSRep2-gDNA								0.997124591346463	0.999898629156486	0.997146313339738	0.999262499336115	0.99631232996164
NSRep2-lexo									1.00153304270205	0.999411135356097	0.996988693792514	1.00010452593214
NSRep3-gDNA										1.0025914426893	1.0001587083859	1.00142208315217
NSRep3-lexo											1.00818511051392	1.00119199953361
NSpool-gDNA												1.00381722517674
NSpool-lexo												

This just takes the correlation from the q<0.001 analysis and divides it by the correlation for the same comparison in the q<0.001 analysis. It is to get an idea of how similar the correlation is for each way of calling peaks.

Window=100kb, stepSize=100kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.95097640131608	1.00006951086612	0.999642047211475	0.9297422081645	-0.30045342372390	0.86280472884382	0.92180915299501	-15.8456071614718	0.89146948913462	0.93773049169219	0.91605335158553
LexoRep1			0.96266595668197	0.968125784126385	1.00229693205992	-1.57976638377185	0.92796727274001	0.96285364547132	-0.67485492106696	0.87423789227241	0.98672199658755	0.89832029829797
LexoRep2				1.00055964588372	0.945361275394575	-0.07811775740946	0.88464773474795	0.91673872978514	132.651123381335	0.89212925870641	0.9583773499583	0.91285453108434
LexoRep3					0.957934088017935	-0.25341752358157	0.89488472882976	0.93168066120805	11.0204471018335	0.89458728591858	0.97538165651346	0.90928994100212
NSRep1-gDNA						0.45017518829032	0.97687279574073	0.73278783566179	0.19908118622136	0.80231333415713	0.96102308046270	1.09878478083973
NSRep1-lexo							1.55880180894535	1.93242574559153	0.65299944274471	1.168825524791	0.48144133907740	0.67919494492504
NSRep2-gDNA								1.30400179910089	0.88326136996044	1.48041123667428	0.98716642478288	1.55013933399494
NSRep2-lexo									1.24902608332192	1.00640909590916	-8.59456170054061	1.03984566975406
NSRep3-gDNA										1.18499559885937	0.70327910495106	1.13148895764153
NSRep3-lexo											3.13881224704161	1.05412833297639
NSpool-gDNA												1.06917274121773
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		0.95964153193654	0.99748715371685	0.996678359773695	0.94397376091465	1.13530574480529	0.84517138251513	0.89879919056371	-2.43330958626334	0.92716812342285	0.86241331013784	1.01101027601912
LexoRep1			0.96217182591319	0.96950166000902	0.97522390919835	0.90527294243390	0.82780783634471	1.0388212839688	-0.27749315385200	1.02622105152044	0.84927340414072	1.18890669215559
LexoRep2				0.994538341304670	0.95038963296744	1.34349931784106	0.86891200186007	0.88203563686105	-2.25015271592607	0.91276603886157	0.87864198080702	0.99031235789166
LexoRep3					0.95802361466750	1.44116705782444	0.83846457672073	0.91870758672094	-7.34138899650364	0.94014931333967	0.87104772678758	1.00828868395558
NSRep1-gDNA						0.59891072232555	1.06964665596083	0.57341519117567	0.53069715296944	0.73970110695701	0.89844783014397	1.34567308411707
NSRep1-lexo							1.34454214564473	1.81162806406634	0.58383626317733	0.68986237718001	0.58337734510220	0.54974005118618
NSRep2-gDNA								1.13214387667467	1.00155452497814	1.208369549477	0.99967908753557	1.21884038478403
NSRep2-lexo									1.21476361618453	1.00955841419775	1.84201034148563	1.06627650868651
NSRep3-gDNA										1.11676625551675	0.92195484964733	1.0911329717889
NSRep3-lexo											1.29718609862852	1.0415399016454
NSpool-gDNA												1.0602641040265
NSpool-lexo												

This just takes the correlation from the q<0.001 analysis and subtracts the correlation for the same comparison in the q<0.001 analysis. It is to get an idea of how similar the correlation is for each way of calling peaks.

Window=100kb, stepSize=100kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool			0.00006844739242	-0.00035054907867	-0.04771599773170	0.04624045566190	-0.05495348867251	0.01916521519993	-0.11986625301882	0.03151918886531	-0.03353283712121	0.02435095602363
LexoRep1			-0.03093225655233	-0.02711656011002	0.00191681242848	0.04936443296092	-0.04673073519144	0.00725170414468	-0.13589107145433	0.03050708210638	-0.00929483138486	0.02472470344263
LexoRep2				0.00054699091791	-0.03704734539713	0.04961256198160	-0.04666800332458	0.02033020849609	-0.1147052346135	0.03133445040447	-0.02224057931055	0.02547614181102
LexoRep3					-0.02914371823709	0.06313873472099	-0.04505547201431	0.01678210157966	-0.11013976598136	0.03118355342135	-0.01341008350210	0.02704132679545
NSRep1-gDNA						-0.12445139217942	-0.01626767236990	0.05706501713593	-0.20618791711190	0.04394269866663	-0.03492571213731	-0.01404506851859
NSRep1-lexo							0.05530950374706	0.10200643875866	-0.13385651556760	0.02781893480851	-0.15493491256564	-0.11570295085405
NSRep2-gDNA								0.09565074456968	-0.05498958735312	0.09884207905186	-0.00999912799197	0.10689303872400
NSRep2-lexo									0.12416723952260	0.00465150460339	0.06910096491095	0.02790293988306
NSRep3-gDNA										0.12675408533014	-0.15928541753731	0.09102587952963
NSRep3-lexo											0.06710171928491	0.04756486958176
NSpool-gDNA												0.00806969560865
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		-0.03274345421460	-0.00244032474989	-0.00321193434449	-0.03903233363606	0.00628497971178	-0.04236997106083	0.02739266302574	-0.09688805421726	0.01992237182498	-0.06162172232034	-0.00271987747028
LexoRep1			-0.03035781607518	-0.02461966912585	-0.01863838046233	-0.01270236095800	-0.06739548504591	-0.00723828075122	-0.12684786317781	-0.00528201765227	-0.07813218820464	-0.03089223276873
LexoRep2				-0.00522810916660	-0.03420834201459	0.01324685279863	-0.03602226167721	0.03131179186213	-0.09055131819174	0.02350626521974	-0.05393741630488	0.00238142365379
LexoRep3					-0.02890963928824	0.01561966404986	-0.04403423317128	0.02219535139728	-0.09393676374006	0.01682451925389	-0.05585992970280	-0.00214648770721
NSRep1-gDNA						-0.15661851757502	0.02591194620062	0.10090957412759	-0.10122629634967	0.04530963662140	-0.07053469848458	-0.01878746295719
NSRep1-lexo							0.03783936207283	0.05884327243801	-0.12032537941979	-0.04711660595852	-0.13023755842858	-0.13893086320627
NSRep2-gDNA								0.06885036572445	0.00087808401311	0.08484016197957	-0.00020634598682	0.08745818269833
NSRep2-lexo									0.10340881816356	0.00701552887582	0.09073480985872	0.04547869572959
NSRep3-gDNA										0.08375679754246	-0.05051979931460	0.06576043010194
NSRep3-lexo											0.07244312140499	0.03625751036571
NSpool-gDNA												0.02121838239833
NSpool-lexo												

						38 negatives	66 total	~57.6% neg				

This just takes the correlation from the q<0.001 analysis and subtracts the correlation for the same comparison in the q<0.001 analysis -- then normalizes it by the q<0.001 correlation to get an idea of how big the difference is relative to the size of the correlation. It is to get an idea of how similar the correlation is for each way of calling peaks.

Window=100kb, stepSize=100kb

Pearson	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		-5.15508046425445	0.00695060346918	-0.03580809646005	-7.55669595491459	432.830289502354	-15.9010801134602	-8.48232486637163	106.310897334572	-12.1743382345849	-6.64044828012755	-9.16394752217959
LexoRep1			-3.87819295560242	-3.29236307887215	0.22916682536331	163.300498749214	-7.76242108703883	-3.85794400876902	248.179996734553	-14.385341660345	-1.34566812722958	-11.3188694382923
LexoRep2				0.05593328553881	-5.77966604170623	1380.11867360497	-13.0393444442507	-9.08233365839904	99.2461428335399	-12.0913802838393	-4.34303357059835	-9.54647930729315
LexoRep3					-4.39131590661951	494.605702820751	-11.746235887576	-7.33291369420025	90.9259579873703	-11.7833905914699	-2.52397031686402	-9.97592241017274
NSRep1-gDNA						-122.135743153195	-2.36747346840874	-36.4651473911115	-402.307635884835	-24.639583741999	-4.05577351155083	8.99036668165672
NSRep1-lexo							35.848162719507	48.2515691854493	-53.1394875004416	14.4440313126449	-107.709625001527	-47.2331335019525
NSRep2-gDNA								23.31298923901	-13.2167707102124	32.4512017183492	-1.30004170471413	35.489669988384
NSRep2-lexo									19.9376207308344	0.63682809855507	111.635264657383	3.83188302966958
NSRep3-gDNA										15.6115009235008	-42.1910579967506	11.6208785559522
NSRep3-lexo											68.1408150187218	5.13489024847265
NSpool-gDNA												6.46974418174471
NSpool-lexo												

Spearman	LexoPool	LexoRep1	LexoRep2	LexoRep3	NSRep1-gDNA	NSRep1-lexo	NSRep2-gDNA	NSRep2-lexo	NSRep3-gDNA	NSRep3-lexo	NSpool-gDNA	NSpool-lexo
LexoPool		-0.04205577470372	-0.00251917658666	-0.00333271029086	-0.05935147925198	0.11918000540770	-0.18319197820460	-0.11259557251360	1.41096291472538	-0.07855304203974	-0.15953683488507	0.01089037003903
LexoRep1			-0.03931540403493	-0.03145774911896	-0.02540554078704	-0.10463922329481	-0.20800982558418	0.03737051268384	4.60369250959366	0.02555107545454	-0.17747711764478	0.15889110003501
LexoRep2				-0.00549165222545	-0.05220002966325	0.25567509657768	-0.15086452697086	-0.13374103971439	1.44441428038294	-0.09557099785091	-0.13811998725751	-0.00978241060119
LexoRep3					-0.04381560609762	0.30611791702373	-0.19265622873543	-0.08848562312324	1.13621400534371	-0.06366083111598	-0.14804271826527	0.00822054644415
NSRep1-gDNA						-0.66969794115059	0.06511183442934	-0.74393705536418	-0.88431385848715	-0.35189739557617	-0.11303068074609	0.25687746020711
NSRep1-lexo							0.25625239547958	0.44801031744042	-0.71280898955783	-0.44956448282880	-0.71415638333505	-0.81904155944667
NSRep2-gDNA								0.11672003832481	0.00155211218098	0.17243859675807	-0.00032101548229	0.17954802574317
NSRep2-lexo									0.17679457412387	0.00946791593564	0.45711488286571	0.06215696223877
NSRep3-gDNA										0.10455747112695	-0.08465181389578	0.08352141686222
NSRep3-lexo											0.22910058852984	0.03988315913751
NSpool-gDNA												0.05683876667863
NSpool-lexo												

Correlation of Genomic Feature densities across genome - Feature v G4

Sample1	Sample2	Pearson	Spearman	Genome partition or overlapping windows?	Window (bin) size	Step size	Details about sample1 peak set	Details about sample2 peak set	Date
Lexo G0 Rep1	G4 motifs	0.8073662	0.6549589	partition	100kb	100kb	p<1e-3, peak set for ASBMB	see G4 data tab	4/2013
NS pool - gDNA control	G4 motifs	0.6814746	0.4328362	partition	100kb	100kb	p<1e-3, peak set for ASBMB	see G4 data tab	4/2013
NS pool - Lexo control	G4 motifs	-0.1931788	-0.2472805	partition	100kb	100kb	p<1e-3, peak set for ASBMB	see G4 data tab	4/2013

Correlation of Genomic Feature densities across genome - Feature v Gene

Sample1	Sample2	Pearson	Spearman	Genome partition or overlapping windows?	Window (bin) size	Step size	Details about sample1 peak set	Details about sample2 peak set	Date
G4 motif	refSeq gene	0.3614393	0.3645467	partition	100kb	100kb	see G4 data tab		4/2013
NS pool - gDNA control	refSeq gene	0.371433	0.2502294	partition	100kb	100kb	p<1e-3, peak set for ASBMB		4/2013
NS pool - Lexo control	refSeq gene	-0.1704692	-0.256417	partition	100kb	100kb	p<1e-3, peak set for ASBMB		4/2013
Dellino ORC	refSeq gene	0.2817503	0.2763272	partition	100kb	100kb			4/2013
Note: above refSeqGenes were not sorted and I did not remove redundant lines with 'uniq'. This will likely not affect the outcome much. Nonetheless, when the analysis is redone, this issue will be resolved.									

Correlation of Genomic Feature densities across genome - Feature v ORC

Sample1	Sample2	Pearson	Spearman	Genome partition or overlapping windows?	Window (bin) size	Step size	Details about sample1 peak set	Details about sample2 peak set	Date
NS pool - gDNA control	Dellino ORC	0.3822271	0.2342207	partition	100kb	100kb	p<1e-3, peak set for ASBMB		4/2013
NS pool - Lexo control	Dellino ORC	-0.1373419	-0.1737509	partition	100kb	100kb	p<1e-3, peak set for ASBMB		4/2013
refSeq gene	Dellino ORC	0.2817503	0.2763272	partition	100kb	100kb			4/2013
G4 motif	Dellino ORC	0.41918	0.3238694	partition	100kb	100kb	see G4 data tab		4/2013

Correlation of Genomic Feature densities across genome - Feature v LexoG0

Sample1	Sample2	Pearson	Spearman	Genome partition or overlapping windows?	Window (bin) size	Step size	Details about sample1 peak set	Details about sample2 peak set	Date
pooled NS gDNA control	Lexo G0 Rep1	0.6924644	0.5070627	partition	100kb	100kb	p<1e-3, peak set for ASBMB	p<1e-3, peak set for ASBMB	4/2013
pooled NS Lexo control	Lexo G0 Rep1	-0.2015242	-0.1557536						4/2013
Dellino ORC	Lexo G0 Rep1	0.4045619	0.3138969						4/2013
G4 motifs	Lexo G0 Rep1	0.8073662	0.6549589	partition	100kb	100kb	p<1e-3, peak set for ASBMB	see G4 data tab	4/2013

For NT% signal: Window=100bp, stepSize=100bp (partition, 100bp resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.564404	0.564412	-0.283681	0.283671	-0.222929	0.222918	-0.480368	0.480375	-0.465215	0.465219	-0.131822	0.13182
LexoRep2	-0.659974	0.659987	-0.419358	0.419363	-0.398068	0.398073	-0.596447	0.59646	-0.585728	0.585738	-0.12853	0.128527
LexoRep3	-0.633225	0.633236	-0.395645	0.395647	-0.374733	0.374736	-0.571866	0.571876	-0.565926	0.565934	-0.129118	0.129115
LexoPool	-0.65546	0.655471	-0.39972	0.399721	-0.381604	0.381605	-0.587394	0.587405	-0.570009	0.570017	-0.127788	0.127785
NS-gDNA pool	-0.37331	0.373321	-0.0934451	0.0934445	-0.0777158	0.0777148	-0.331176	0.331186	-0.353551	0.353559	-0.127788	0.127785
NS-lexo pool	-0.160446	0.16046	-0.000531187	0.000533943	0.00913602	-0.0091336	0.17703	-0.177008	-0.354004	0.354012	-0.462345	0.462334
NSRep1-gDNA	-0.531025	0.531032	-0.234607	0.234603	-0.172507	0.172502	-0.45831	0.458316	-0.460452	0.460456	-0.132376	0.132374
NSRep1-lexo	-0.403904	0.403917	-0.169661	0.169661	-0.0724232	0.0724205	-0.0325993	0.0326186	-0.460452	0.460456	-0.463753	0.463743
NSRep2-gDNA	-0.354363	0.354377	-0.130099	0.130101	-0.0936166	0.0936179	-0.312257	0.31227	-0.336379	0.336389	-0.130937	0.130935
NSRep2-lexo	-0.101108	0.101123	0.00808075	-0.00807725	0.0195828	-0.0195806	0.233964	-0.233939	-0.336379	0.336389	-0.463314	0.463304
NSRep3-gDNA	-0.207059	0.207067	-0.049188	0.0491889	-0.0316874	0.0316879	-0.171233	0.171241	-0.224196	0.224202	-0.131101	0.131099
NSRep3-lexo	-0.0712403	0.0712497	0.00421535	-0.00421207	0.0142685	-0.0142656	0.210068	-0.210053	-0.224196	0.224202	-0.463404	0.463394

For NT% signal: Window=100bp, stepSize=50bp (overlapping windows, 50bp resolution, Score for 100bp window assign nt 25-74 within it)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.566755	0.566762	-0.284608	0.284598	-0.223562	0.22355	-0.482231	0.482238	-0.466936	0.46694	-0.132146	0.132144
LexoRep2	-0.662832	0.662845	-0.420369	0.420374	-0.398908	0.398912	-0.598846	0.598859	-0.58799	0.588001	-0.128836	0.128834
LexoRep3	-0.635958	0.635969	-0.396601	0.396604	-0.375534	0.375536	-0.574162	0.574173	-0.568109	0.568117	-0.129424	0.129421
LexoPool	-0.658275	0.658286	-0.400678	0.400679	-0.382426	0.382427	-0.58974	0.589752	-0.572194	0.572202	-0.128088	0.128085
NS-gDNA pool	-0.375422	0.375433	-0.0940024	0.0940018	-0.0781765	0.0781756	-0.332966	0.332977	-0.355268	0.355276	-0.128088	0.128085
NS-lexo pool	-0.162259	0.162272	-0.00107273	0.00107551	0.00875297	-0.00875053	0.176036	-0.176015	-0.355725	0.355733	-0.463056	0.463045
NSRep1-gDNA	-0.534434	0.53444	-0.23627	0.236266	-0.173762	0.173757	-0.461117	0.461122	-0.463061	0.463065	-0.132709	0.132707
NSRep1-lexo	-0.407804	0.407817	-0.172745	0.172746	-0.0742301	0.0742276	-0.0352733	0.0352928	-0.463061	0.463065	-0.464513	0.464503
NSRep2-gDNA	-0.355955	0.355968	-0.130499	0.130501	-0.0938377	0.093839	-0.313535	0.313548	-0.337642	0.337652	-0.131252	0.13125
NSRep2-lexo	-0.102437	0.102452	0.00771167	-0.00770816	0.0194456	-0.0194434	0.233551	-0.233526	-0.337642	0.337652	-0.464054	0.464044
NSRep3-gDNA	-0.208307	0.208315	-0.0496146	0.0496155	-0.032003	0.0320035	-0.172274	0.172282	-0.225341	0.225347	-0.131424	0.131422
NSRep3-lexo	-0.0721348	0.0721443	0.00388021	-0.00387693	0.014048	-0.0140451	0.209701	-0.209686	-0.225341	0.225347	-0.464154	0.464144

For NT% signal: Window=10bp, stepSize=10bp (partition, 10bp resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.341748	0.34175	-0.170985	0.170978	-0.134078	0.13407	-0.290502	0.290503	-0.280835	0.280835	-0.0787058	0.0787036
LexoRep2	-0.40027	0.400274	-0.251921	0.251922	-0.238761	0.238761	-0.361225	0.361229	-0.354214	0.354217	-0.0767305	0.0767281
LexoRep3	-0.384033	0.384036	-0.237606	0.237605	-0.22471	0.224709	-0.346346	0.346349	-0.342248	0.34225	-0.0770787	0.0770764
LexoPool	-0.39743	0.397433	-0.240076	0.240074	-0.228905	0.228903	-0.35566	0.355663	-0.34462	0.344622	-0.076282	0.0762796
NS-gDNA pool	-0.228201	0.228206	-0.0572537	0.0572529	-0.0476276	0.0476267	-0.202212	0.202217	-0.214939	0.214941	-0.076282	0.0762795
NS-lexo pool	-0.101971	0.101979	-0.00260032	0.00260203	0.00372628	-0.00372473	0.100589	-0.100575	-0.215215	0.215218	-0.275386	0.275377
NSRep1-gDNA	-0.324919	0.32492	-0.143958	0.143955	-0.105965	0.105962	-0.279968	0.279968	-0.280416	0.280416	-0.0790419	0.0790401
NSRep1-lexo	-0.252133	0.252138	-0.110816	0.110815	-0.0489445	0.0489433	-0.0282501	0.0282618	-0.280416	0.280416	-0.27629	0.276281
NSRep2-gDNA	-0.215637	0.215643	-0.0785198	0.0785205	-0.0562916	0.0562919	-0.189728	0.189734	-0.203585	0.20359	-0.0781708	0.078169
NSRep2-lexo	-0.0655833	0.065592	0.00296784	-0.00296569	0.0107922	-0.0107908	0.135757	-0.135741	-0.203585	0.20359	-0.276002	0.275994
NSRep3-gDNA	-0.127116	0.127119	-0.0307221	0.0307224	-0.0199742	0.0199743	-0.105321	0.105325	-0.136761	0.136764	-0.0782736	0.0782716
NSRep3-lexo	-0.0460504	0.0460556	0.000928629	-0.000926635	0.0073545	-0.00735272	0.122195	-0.122185	-0.136761	0.136764	-0.276066	0.276057

For NT% signal: Window=200bp, stepSize=200bp (partition, 200bp resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.619164	0.619175	-0.313209	0.313198	-0.246782	0.246768	-0.528145	0.528153	-0.511759	0.511764	-0.145417	0.145414
LexoRep2	-0.722824	0.722841	-0.463776	0.463783	-0.440865	0.440872	-0.65434	0.654356	-0.643028	0.643041	-0.141933	0.14193
LexoRep3	-0.693779	0.693793	-0.437517	0.437521	-0.414932	0.414936	-0.627574	0.627587	-0.621458	0.621468	-0.142561	0.142558
LexoPool	-0.71816	0.718175	-0.442207	0.44221	-0.422654	0.422656	-0.644662	0.644676	-0.626026	0.626036	-0.141141	0.141137
NS-gDNA pool	-0.405173	0.405186	-0.1014	0.101399	-0.0843767	0.0843758	-0.360258	0.36027	-0.385795	0.385804	-0.14114	0.141137
NS-lexo pool	-0.169497	0.169513	0.00214745	-0.00214448	0.0120699	-0.0120673	0.203358	-0.203334	-0.38629	0.386299	-0.512665	0.512655
NSRep1-gDNA	-0.574832	0.57484	-0.253605	0.253601	-0.186443	0.186437	-0.497312	0.497319	-0.5008	0.500805	-0.145994	0.145992
NSRep1-lexo	-0.430863	0.430879	-0.1742	0.174201	-0.0723538	0.0723508	-0.0223586	0.0223805	-0.5008	0.500805	-0.513832	0.513821
NSRep2-gDNA	-0.3873	0.387317	-0.143514	0.143517	-0.103715	0.103717	-0.342331	0.342347	-0.369365	0.369377	-0.144489	0.144487
NSRep2-lexo	-0.105961	0.105979	0.0109058	-0.0109019	0.0225664	-0.022564	0.263328	-0.263301	-0.369365	0.369377	-0.51353	0.513519
NSRep3-gDNA	-0.223665	0.223675	-0.0522485	0.0522496	-0.0333853	0.0333859	-0.185024	0.185033	-0.243729	0.243736	-0.144669	0.144666
NSRep3-lexo	-0.0743649	0.0743755	0.00670823	-0.00670466	0.0172364	-0.0172333	0.237238	-0.237222	-0.243729	0.243736	-0.513609	0.513599

Note: None of the above have ChrY or ChrM removed. Below I repeated the 200bp partition after removing chrY and chrM entries for the NT scores. Though the signal files still have chrM and chrY, the correlation script will not compare them as it only compares bp that are covered in both files. I also checked just once to make sure this was true (i.e. rmeoved chrY and chrM entries from one of the sample FE signal files and did the correlation). The result was precisely the same.

For NT% signal: Window=200bp, stepSize=200bp (partition, 200bp resolution, chrM and chrY entries removed)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.642384	0.642395	-0.37909	0.379077	-0.310125	0.310108	-0.542646	0.542654	-0.527025	0.52703	-0.146904	0.146902
LexoRep2	-0.727402	0.727419	-0.46927	0.469278	-0.446325	0.446332	-0.657376	0.657393	-0.647926	0.64794	-0.143462	0.143459
LexoRep3	-0.698724	0.698739	-0.443352	0.443356	-0.420707	0.420711	-0.630891	0.630905	-0.626712	0.626723	-0.144083	0.14408
LexoPool	-0.725252	0.725268	-0.45458	0.454583	-0.435084	0.435087	-0.649234	0.649249	-0.632833	0.632844	-0.142673	0.14267
NS-gDNA pool	-0.440225	0.440239	-0.190521	0.19052	-0.16568	0.165678	-0.381674	0.381687	-0.407551	0.40756	-0.142673	0.14267
NS-lexo pool	-0.169595	0.169611	0.00199083	-0.00198774	0.0124276	-0.0124249	0.209629	-0.209605	-0.408129	0.408139	-0.519156	0.519145
NSRep1-gDNA	-0.614612	0.614621	-0.37228	0.372274	-0.295248	0.295239	-0.519266	0.519273	-0.522684	0.522689	-0.147464	0.147462
NSRep1-lexo	-0.431608	0.431624	-0.188971	0.188972	-0.0837509	0.0837474	-0.0216897	0.0217123	-0.522684	0.522689	-0.520188	0.520177
NSRep2-gDNA	-0.41976	0.419779	-0.249956	0.249961	-0.201034	0.201038	-0.360427	0.360444	-0.387404	0.387416	-0.145991	0.145989
NSRep2-lexo	-0.105613	0.105631	0.0110599	-0.011056	0.0232572	-0.0232548	0.269411	-0.269383	-0.387404	0.387416	-0.519919	0.519908
NSRep3-gDNA	-0.237557	0.237567	-0.0749786	0.07498	-0.0502772	0.050278	-0.193425	0.193435	-0.255915	0.255922	-0.146173	0.14617
NSRep3-lexo	-0.074144	0.0741547	0.0066692	-0.00666557	0.0173679	-0.0173648	0.242625	-0.242609	-0.255915	0.255922	-0.519986	0.519976

For NT% signal: Window=1kb, stepSize=1kb (partition, 1kb resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.651365	0.651373	-0.349646	0.349612	-0.280265	0.280227	-0.565284	0.565278	-0.551209	0.5512	-0.162116	0.162109
LexoRep2	-0.751373	0.751392	-0.511612	0.511616	-0.48911	0.489114	-0.688099	0.688113	-0.680846	0.680856	-0.158964	0.158957
LexoRep3	-0.723742	0.723758	-0.481995	0.481991	-0.459408	0.459403	-0.662187	0.662194	-0.659946	0.65995	-0.1596	0.159593
LexoPool	-0.749143	0.749159	-0.489906	0.489899	-0.470384	0.470376	-0.680484	0.680493	-0.665339	0.665343	-0.158255	0.158248
NS-gDNA pool	-0.386076	0.386085	-0.100907	0.100896	-0.0846329	0.0846208	-0.349403	0.349406	-0.385959	0.385958	-0.158255	0.158248
NS-lexo pool	-0.108879	0.108896	0.0217414	-0.0217388	0.0267425	-0.0267402	0.293592	-0.293567	-0.386443	0.386442	-0.585264	0.585242
NSRep1-gDNA	-0.551245	0.551247	-0.249957	0.249929	-0.185673	0.185642	-0.485642	0.485636	-0.498815	0.498806	-0.162613	0.162606
NSRep1-lexo	-0.351956	0.351973	-0.117293	0.117285	-0.0464944	0.0464816	0.0775205	-0.077505	-0.498815	0.498806	-0.583913	0.583892
NSRep2-gDNA	-0.389952	0.389965	-0.159304	0.159298	-0.119135	0.119129	-0.353327	0.353336	-0.388518	0.388522	-0.161272	0.161267
NSRep2-lexo	-0.0544219	0.0544405	0.0311697	-0.0311658	0.0335144	-0.0335119	0.336703	-0.33667	-0.388518	0.388522	-0.584769	0.584748
NSRep3-gDNA	-0.195495	0.195503	-0.0407429	0.0407378	-0.0243809	0.0243753	-0.160047	0.160051	-0.228461	0.228462	-0.161444	0.161437
NSRep3-lexo	-0.0287253	0.0287368	0.0257316	-0.0257278	0.0308571	-0.0308538	0.315445	-0.315427	-0.228461	0.228462	-0.584745	0.584725

For NT% signal: Window=500bp, stepSize=500bp (partition, 500bp resolution)

	log2(Fold Enrichment) Signal		Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.516748	0.516788	-0.66089	0.660902	-0.343461	0.343438	-0.273111	0.273084	-0.568219	0.56822	-0.551844	0.551842	-0.158816	0.15881
LexoRep2	-0.599644	0.599688	-0.768607	0.768628	-0.507973	0.50798	-0.484503	0.48451	-0.69957	0.699588	-0.689168	0.68918	-0.155316	0.15531
LexoRep3	-0.602312	0.602355	-0.739431	0.739449	-0.479076	0.479079	-0.455679	0.455681	-0.672372	0.672385	-0.667293	0.667301	-0.155974	0.155967
LexoPool	-0.591394	0.591444	-0.764851	0.764869	-0.485242	0.485242	-0.465039	0.465038	-0.690385	0.690398	-0.672063	0.672072	-0.154535	0.154529
NS-gDNA pool	-0.495402	0.495517	-0.413435	0.413448	-0.105356	0.10535	-0.0881222	0.088116	-0.37071	0.370719	-0.402086	0.402091	-0.154535	0.154529
NS-lexo pool	-	-	-0.153201	0.153218	0.0102497	-0.0102465	0.0182711	-0.0182684	0.248651	-0.248623	-0.402596	0.4026	-0.56387	0.563852
NSRep1-gDNA	-0.520675	0.520708	-0.587724	0.58773	-0.261617	0.2616	-0.1933	0.193281	-0.512714	0.512714	-0.520594	0.520593	-0.159368	0.159363
NSRep1-lexo	-0.472621	0.472654	-0.417567	0.417586	-0.155014	0.155011	-0.0624249	0.0624175	0.0175012	-0.0174787	-0.520594	0.520593	-0.563687	0.563669
NSRep2-gDNA	-	-	-0.404865	0.404882	-0.156406	0.156405	-0.114968	0.114966	-0.361817	0.361831	-0.393458	0.393467	-0.157873	0.157868
NSRep2-lexo	-	-	-0.0915369	0.0915566	0.0188238	-0.0188196	0.0273208	-0.0273183	0.303146	-0.303112	-0.393458	0.393467	-0.564005	0.563987
NSRep3-gDNA	-	-	-0.220457	0.220466	-0.0493136	0.0493118	-0.0309792	0.0309769	-0.18207	0.182076	-0.247447	0.247451	-0.158074	0.158069
NSRep3-lexo	-	-	-0.0596854	0.0596972	0.0148662	-0.0148623	0.0234508	-0.0234474	0.279183	-0.279164	-0.247447	0.247451	-0.564027	0.56401

*Note: tried log2(FE) here with 500bp windows --- tried log10(FE) below for 1kb.

For NT% signal: Window=5kb, stepSize=5kb (partition, 5kb resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.581958	0.581952	-0.371214	0.371173	-0.308713	0.308668	-0.507963	0.507945	-0.507136	0.507114	-0.161933	0.16192
LexoRep2	-0.640322	0.64033	-0.440377	0.440376	-0.420571	0.42057	-0.591499	0.591499	-0.597651	0.597645	-0.15991	0.159897
LexoRep3	-0.616708	0.616713	-0.415313	0.415307	-0.395452	0.395446	-0.569661	0.569657	-0.57977	0.579761	-0.160427	0.160414
LexoPool	-0.642499	0.642504	-0.430363	0.430355	-0.413249	0.413241	-0.588224	0.588219	-0.5882	0.58819	-0.159449	0.159436
NS-gDNA pool	-0.338719	0.338718	-0.144463	0.144441	-0.124012	0.123989	-0.303167	0.303162	-0.351829	0.351818	-0.159449	0.159436
NS-lexo pool	-0.028342	0.0283497	0.0461942	-0.0461918	0.0463764	-0.0463742	0.360566	-0.360549	-0.352352	0.352341	-0.588171	0.588148
NSRep1-gDNA	-0.493089	0.493083	-0.306816	0.30678	-0.243698	0.243657	-0.431511	0.431497	-0.455632	0.455614	-0.16198	0.161969
NSRep1-lexo	-0.239152	0.239159	-0.0627734	0.0627664	-0.0247931	0.0247824	0.171535	-0.171531	-0.455632	0.455614	-0.585322	0.585301
NSRep2-gDNA	-0.366399	0.3664	-0.257365	0.257354	-0.217824	0.217811	-0.33148	0.33148	-0.374831	0.374825	-0.161508	0.161497

NSRep2-lexo	0.0116218	-0.0116148	0.0551555	-0.0551534	0.0466358	-0.0466336	0.376663	-0.376642	-0.374831	0.374825	-0.586858	0.586837		
NSRep3-gDNA	-0.139164	0.139163	-0.0165361	0.0165269	0.000835649	-0.000845788	-0.105043	0.105041	-0.180956	0.180948	-0.161519	0.161506		
NSRep3-lexo	0.0323126	-0.0323069	0.0504738	-0.0504707	0.0493198	-0.0493168	0.371088	-0.371077	-0.180956	0.180948	-0.586735	0.586715		

For NT% signal: Window=1.5kb, stepSize=1.5kb (partition, 1.5kb resolution)

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal			
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC		
LexoRep1	-0.658381	0.65838	-0.417571	0.41751	-0.348396	0.348326	-0.568675	0.568652	-0.557895	0.55787	-0.164049	0.164036		
LexoRep2	-0.732134	0.732149	-0.505623	0.505624	-0.483814	0.483813	-0.671516	0.671525	-0.668851	0.668853	-0.16119	0.161177		
LexoRep3	-0.705832	0.705843	-0.476785	0.476775	-0.454831	0.45482	-0.646793	0.646794	-0.648967	0.648962	-0.161783	0.16177		
LexoPool	-0.733371	0.733382	-0.492617	0.492603	-0.473771	0.473756	-0.66653	0.66653	-0.656582	0.656576	-0.160529	0.160516		
NS-gDNA pool	-0.397211	0.397217	-0.180511	0.180487	-0.158096	0.15807	-0.352075	0.352073	-0.39283	0.392821	-0.160529	0.160516		
NS-lexo pool	-0.0819229	0.0819355	0.0293034	-0.0293014	0.0328943	-0.0328926	0.325315	-0.325288	-0.393383	0.393374	-0.5979	0.59787		
NSRep1-gDNA	-0.561434	0.561432	-0.351921	0.351879	-0.28232	0.282272	-0.485576	0.485564	-0.502305	0.502289	-0.164405	0.164393		
NSRep1-lexo	-0.313653	0.313667	-0.104507	0.104498	-0.0438095	0.0437954	0.115204	-0.115186	-0.502305	0.502289	-0.595856	0.595827		
NSRep2-gDNA	-0.407157	0.407165	-0.27533	0.275318	-0.231116	0.231101	-0.36128	0.361284	-0.399095	0.399093	-0.163311	0.163301		
NSRep2-lexo	-0.0302455	0.0302601	0.0394996	-0.0394964	0.0381913	-0.0381892	0.360597	-0.36056	-0.399095	0.399093	-0.597014	0.596985		
NSRep3-gDNA	-0.189506	0.189509	-0.0493259	0.0493165	-0.0290813	0.0290711	-0.150948	0.150947	-0.224721	0.224714	-0.163447	0.163433		
NSRep3-lexo	-0.00976564	0.00977395	0.0324701	-0.0324669	0.0356468	-0.0356438	0.340383	-0.340364	-0.224721	0.224714	-0.596941	0.596913		

For NT% signal: Window=1kb, stepSize=1kb (partition, 1kb resolution, chrM and chrY removed)

	log10(Fold Enrichment) Signal		Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1	-0.471764	0.471806	-0.6756	0.675608	-0.423043	0.423002	-0.35213	0.352082	-0.580558	0.580551	-0.567345	0.567336	-0.163648	0.163641
LexoRep2	-0.53346	0.533506	-0.755995	0.756015	-0.517666	0.517671	-0.495168	0.495171	-0.691183	0.691198	-0.685848	0.685858	-0.160554	0.160547
LexoRep3	-0.539752	0.539797	-0.728764	0.72878	-0.488405	0.488402	-0.465791	0.465786	-0.665572	0.66558	-0.665338	0.665342	-0.16118	0.161173
LexoPool	-0.518966	0.519019	-0.756396	0.756412	-0.503588	0.503581	-0.484198	0.484191	-0.68518	0.685188	-0.672364	0.672368	-0.159852	0.159845
NS-gDNA pool	-0.441515	0.441657	-0.419285	0.419296	-0.189815	0.189793	-0.166473	0.166449	-0.369927	0.369931	-0.407414	0.407413	-0.159852	0.159845
NS-lexo pool	-0.253963	0.254133	-0.108393	0.10841	0.0225061	-0.0225034	0.0278255	-0.0278231	0.302174	-0.302148	-0.407981	0.40798	-0.592535	0.592514
NSRep1-gDNA	-0.472548	0.472581	-0.589137	0.589139	-0.366917	0.366875	-0.294197	0.294148	-0.506767	0.50676	-0.520235	0.520226	-0.164119	0.164112
NSRep1-lexo	-0.427387	0.427421	-0.352037	0.352055	-0.126932	0.126923	-0.0538156	0.0538006	0.0816372	-0.0816213	-0.520235	0.520226	-0.590991	0.59097
NSRep2-gDNA	-0.294153	0.294236	-0.42244	0.422454	-0.27727	0.27726	-0.230825	0.230813	-0.371805	0.371814	-0.407224	0.407229	-0.162824	0.162819
NSRep2-lexo	-0.16725	0.167338	-0.0534771	0.0534959	0.0320113	-0.0320073	0.034638	-0.0346354	0.344245	-0.344212	-0.407224	0.407229	-0.591902	0.591881
NSRep3-gDNA	-0.309403	0.309463	-0.207603	0.20761	-0.0588758	0.0588685	-0.0372687	0.0372602	-0.167191	0.167195	-0.239678	0.239678	-0.162994	0.162988
NSRep3-lexo	-0.224739	0.224802	-0.0281605	0.0281721	0.0260199	-0.026016	0.0312098	-0.0312065	0.322303	-0.322285	-0.239678	0.239678	-0.591867	0.591847

*Note: tried log10(FE) here with 1kb windows --- ran a pilot for log2(FE) above for 500bp.

...

	Fold Enrichment Signal		-log10(pval) Signal		-log10(qval) Signal		Subtraction Signal		Treatment Pileup Signal		Control Lambda Signal	
	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC	percent AT	percent GC
LexoRep1												
LexoRep2												
LexoRep3												
LexoPool												
NS-gDNA pool												
NS-lexo pool												
NSRep1-gDNA												
NSRep1-lexo												
NSRep2-gDNA												
NSRep2-lexo												
NSRep3-gDNA												
NSRep3-lexo												

Peak and Feature Counts in 100kb bins, gaps subtracted, and keeping only 100kb bins thereafter -- chrY and chrM removed.

Date: 8/6/2013	G4		CpG Islands		Genes	
q < 0.001 peak sets	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
LexoPool	0.704340510230235	0.704259182270841	0.479103534608883	0.500701954040099	0.316339173500489	0.328808873445982
LexoRep1	0.809706097912294	0.644004285762933	0.617389472013765	0.57288971666726	0.369182537003929	0.359890377895182
LexoRep2	0.711688943479151	0.697370107104918	0.483775643676568	0.493376817936936	0.316747061948343	0.322558609372032
LexoRep3	0.734645230528268	0.70827825795823	0.502227499056266	0.50852458710213	0.324738565134453	0.332789689919857
NSpool-gDNA	0.692298958861982	0.362984758861691	0.627342305012026	0.355865870177047	0.387078760683999	0.188723063064132
NSRep1-gDNA	0.788719935750759	0.541384354701493	0.697885163241926	0.53588197191013	0.439835131746807	0.359622206849903
NSRep2-gDNA	0.584887124736971	0.203652501788213	0.464809643068689	0.115322856210294	0.21481433428132	0.013234477876023
NSRep3-gDNA	-0.073731038283802	-0.100230102989417	-0.038872708019614	-0.128386647664135	-0.093567292737904	-0.174592324259717
NSpool-lexo	-0.248159204213938	-0.259855638692166	-0.215430740524686	-0.29113801666377	-0.211611963995751	-0.277969472800367
NSRep1-lexo	0.023751908826318	0.038828730584539	0.086406374944882	0.067563988972945	0.01681901916254	0.031606033990946
NSRep2-lexo	-0.213850677829809	-0.255317174569095	-0.209624795746903	-0.325551401848117	-0.205854717471301	-0.290983597596024
NSRep3-lexo	-0.251991247095849	-0.280384057000954	-0.236058674102866	-0.339721478214607	-0.238128447026328	-0.326140117902255

Date: 8/8/2013	G4		CpG Islands		Genes	
p < 0.001 peak sets	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
LexoPool	0.673070002663163	0.70571962647582	0.456143595041595	0.486821663100458	0.310283995221824	0.324676893715525
LexoRep1	0.815062033673382	0.666406031162124	0.61149275809232	0.55485946905628	0.373306569650802	0.353970408926774
LexoRep2	0.677171192820274	0.701001539686439	0.455244166351404	0.479399467382965	0.30629677420757	0.316606618500899
LexoRep3	0.698466091946205	0.709701427215171	0.470677076893863	0.488362908536878	0.311002733995774	0.323805447868904
NSpool-gDNA	0.688263803304113	0.429961385093827	0.629386831719798	0.434161483991484	0.418428774333294	0.26110086684887
NSRep1-gDNA	0.797799154089746	0.593522670134796	0.697973935498698	0.588336260391764	0.468908713475498	0.417583781232628
NSRep2-gDNA	0.645678161993204	0.266192284263389	0.519305661990166	0.179710182470209	0.265138343862439	0.047660275226174
NSRep3-gDNA	0.08760941684567	0.015440533078805	0.149387943881703	0.025678421811334	0.06711492133821 4	-0.049450266149066
NSpool-lexo	-0.258138026076337	-0.241597194870908	-0.183166854124888	-0.217079357902375	-0.163717889412377	-0.212400324915169
NSRep1-lexo	-0.009938604560448	0.051066700757424	0.108087290026188	0.161389908018393	0.135240254204678	0.153308960133277
NSRep2-lexo	-0.217128983055742	-0.269157441650242	-0.208803937870896	-0.334814476717 38	-0.212528640351769	-0.306804255227955
NSRep3-lexo	-0.27483333868574	-0.288390091038647	-0.236518377608371	-0.313403583234114	-0.2318796316397	-0.304262360280515

Peak Counts in 100kb bins, gaps subtracted, and keeping only 100kb bins thereafter -- chrY and chrM removed.

Date: 8/6/2013	ORC (Dellino)		Bubble-Trap Seq (Mesner)		Aladjem K562 15% FDR set		Aladjem MCF7 15% FDR set	
q < 0.001 peak sets	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
LexoPool	0.359833780326581	0.321613533459584	0.299059426129326	0.237238880929338	0.692967561858443	0.636321323400401	0.741364774206703	0.65672160550199
LexoRep1	0.405289389886313	0.310707124879334	0.258198680511049	0.194197471581593	0.664296644188677	0.527127750948224	0.801531401729978	0.566913543565049
LexoRep2	0.360064908864355	0.318603708984143	0.303042624059054	0.240737635786254	0.701195991495778	0.635251094986827	0.753207087501403	0.660763733672361
LexoRep3	0.367136565218594	0.323887037439276	0.303184085027192	0.239218550431979	0.712575187290772	0.637098758042083	0.769195787242777	0.659094397686491
NSpool-gDNA	0.366125971271429	0.183826002135625	0.219716762714987	0.098919306286607	0.421152975955362	0.198068094522327	0.515913906180161	0.215641973587193
NSRep1-gDNA	0.431571501249422	0.299385297499888	0.273907071283482	0.183868053018164	0.548506926104269	0.401592073447371	0.678424778232832	0.443486452671962
NSRep2-gDNA	0.242572322153951	0.05220510589875	0.086962431528365	-0.020602171939328	0.354268762160354	0.095991926552738	0.447138719511876	0.097180116067225
NSRep3-gDNA	-0.068429926727364	-0.108498814213841	-0.116810080362663	-0.134602385761142	-0.156354972945733	-0.20665533486054	-0.177751385492954	-0.224507384275456
NSpool-lexo	-0.174737747629397	-0.210341402993008	-0.178037179999728	-0.206161579708272	-0.250574171474612	-0.331604493608481	-0.286190008224745	-0.355089030410816
NSRep1-lexo	0.037367979998503	0.030105565883578	0.003943849969416	-0.007078744225597	-0.012255505412326	-0.01904522048775	-0.016054958696183	-0.012946067123398
NSRep2-lexo	-0.16591244983419	-0.217980401676231	-0.184271037882858	-0.224798784069641	-0.204248929985771	-0.290411832502475	-0.226864816643744	-0.303666500304253
NSRep3-lexo	-0.188762687911201	-0.234264999700566	-0.191173584879375	-0.22583915152051	-0.243497714345214	-0.327616592422385	-0.277626537185578	-0.357204257965182

Date: 8/8/2013	ORC (Dellino)		Bubble-Trap Seq (Mesner)		Aladjem K562 15% FDR set		Aladjem MCF7 15% FDR set	
p < 0.001 peak sets	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
LexoPool	0.353260800872584	0.323072605181081	0.300090478956535	0.239400819281097	0.669394698806543	0.641238313944201	0.706332854069495	0.656726113482948
LexoRep1	0.407456927997098	0.315205440250502	0.273084039813366	0.199067787560612	0.694499535675835	0.551540367559257	0.807788896538463	0.585055693795025
LexoRep2	0.351670951616731	0.318600551998739	0.302621150842258	0.243361769946121	0.675216692934466	0.643232986246509	0.714851015666365	0.660465169618993
LexoRep3	0.358207498693384	0.320893125354978	0.303185710333364	0.239388957332406	0.692926576231605	0.647424174876831	0.736379540052278	0.658891727659329
NSpool-gDNA	0.384569508004575	0.232735866112445	0.249614455877508	0.136304439754017	0.432382093222246	0.250066982196407	0.518018568472418	0.268753418159816
NSRep1-gDNA	0.44818704595575	0.336289682816578	0.306663917970967	0.209856947917871	0.564086107295554	0.445051532678867	0.681776722142476	0.476631482597062
NSRep2-gDNA	0.286908563655187	0.089630658357924	0.134982795599056	0.017885560803877	0.419000959896879	0.148191525987038	0.513417310674962	0.158979101927235
NSRep3-gDNA	0.054894506808422	-0.020115000665295	-0.018235541633647	-0.05819447353715	-0.07536555354919	-0.126067157965233	-0.078127448726437	-0.133093562929198
NSpool-lexo	-0.144299235425409	-0.166143100925818	-0.157816919971488	-0.184056261829014	-0.279264625393838	-0.343160362765021	-0.313546732360026	-0.351585537293572
NSRep1-lexo	0.075260941059997	0.066887601449439	0.036211231801159	-0.002896732463649	-0.094064367279292	-0.074896562216593	-0.089706883517499	-0.051865744196722
NSRep2-lexo	-0.168045844138608	-0.226490223460085	-0.181323991218098	-0.225624288763467	-0.209127052990095	-0.303979055710991	-0.232191669303845	-0.319401170281904
NSRep3-lexo	-0.185336603118434	-0.221415783060873	-0.189370555067901	-0.217988492614321	-0.271070669785825	-0.349983684135669	-0.308508333643089	-0.371713676970008

Peak and Feature Counts in 100kb bins, gaps subtracted, and keeping only 100kb bins thereafter -- chrY and chrM removed.														
Date: 8/6/2013	G4		CpG Islands		Genes		ORC (Dellino)		Bubble-Trap Seq (Mesner)		Aladjem K562 15% FDR set		Aladjem MCF7 15% FDR set	
	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
G4	1	1	0.717449505565544	0.551877627547619	0.410970999328038	0.374825386745002	0.423201211535344	0.322453089816646	0.26501775293178	0.228628229273183	0.659931635572556	0.576775418575122	0.733702689023046	0.562572491526797
CpG Islands	0.717449505565544	0.551877627547619	1	1	0.511815740467245	0.561445851414835	0.413759736586892	0.360726709997884	0.24454404697497	0.227873961078101	0.432573020492511	0.391320732522855	0.49859903851743	0.405870007429787
Genes	0.410970999328038	0.374825386745002	0.511815740467245	0.561445851414835	1	1	0.318163121007384	0.277819553183403	0.201484381240694	0.138541850274835	0.278683966254428	0.279346986232637	0.306141168287836	0.291779297640224
ORC (Dellino)	0.423201211535344	0.322453089816646	0.413759736586892	0.360726709997884	0.318163121007384	0.277819553183403	1	1	0.227434445159224	0.187507421257577	0.308340945155608	0.273843299842933	0.3748095505503	0.305303577282765
BubbleTrap-seq (Mesner, pre-pub)	0.26501775293178	0.228628229273183	0.24454404697497	0.227873961078101	0.201484381240694	0.138541850274835	0.227434445159224	0.187507421257577	1	1	0.247354461196706	0.237302099520311	0.269917823058962	0.248491228683186
Aladjem K562	0.659931635572556	0.576775418575122	0.432573020492511	0.391320732522855	0.278683966254428	0.279346986232637	0.308340945155608	0.273843299842933	0.247354461196706	0.237302099520311	1	1	0.718024926235767	0.642993483042698
Aladjem MCF7	0.733702689023046	0.562572491526797	0.49859903851743	0.405870007429787	0.306141168287836	0.291779297640224	0.3748095505503	0.305303577282765	0.269917823058962	0.248491228683186	0.718024926235767	0.642993483042698	1	1

Overlap_analysis_of_replicates_q1e-3sets_noChrYchrM - LexoG0

Overlap of q<0.001 Lexo G0 peak sets

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportion	fractionOfMax
LGR1	LGR1	7458.25065441778	110704	0.0673711036134	1	0	14.8431589563735	1	1	1
LGR1	LGR2	15252.3089482641	102960	0.137775590297225	0.930047694753577	0	6.75045334770238	0.570565648756604	1	0.930047694753577
LGR1	LGR3	14156.8444797592	101496	0.127880153199154	0.916823240352652	0	7.16939429158203	0.602890721155417	1	0.916823240352652
LGR1	LGPool	15743.642145067	105998	0.142213850855136	0.95749024425495	0	6.73274957746755	0.562374587886269	1	0.95749024425495
LGR2	LGR1	15369.5476439576	79055	0.079214264367776	0.407447493879655	0	5.14361267041451	1.75264669749964	0.570565648756604	0.714111504552681
LGR2	LGR2	30788.6391853734	194025	0.158683876744612	1	0	6.30183746776877	1	1	1
LGR2	LGR3	28647.4271536346	159841	0.147648123456434	0.823816518489885	0	5.57959355801069	1.05665443138622	0.946383198041489	0.870489374911503
LGR2	LGPool	31712.4858985705	178511	0.163445359611238	0.920041231800026	0	5.62904467883569	0.98564396421659	1	0.920041231800026
LGR3	LGR1	14251.9850004457	80497	0.077615890255229	0.438384289464225	0	5.64812550655103	1.65867538661656	0.602890721155417	0.727137230813701
LGR3	LGR2	28619.9603238136	165512	0.155863460390441	0.90137347376676	0	5.78309676629019	0.94638319804149	1	0.90137347376676
LGR3	LGR3	26621.7658668381	183622	0.144981352271722	1	0	6.89743876940682	1	1	1
LGR3	LGPool	29486.2988119775	173809	0.160581514262874	0.946558691224363	0	5.89456822330641	0.932796886985588	1	0.946558691224363
LGPool	LGR1	15868.7946415307	79185	0.080613228490232	0.402258561043632	0	4.98998202376144	1.77817423038011	0.562374587886269	0.715285807197572
LGPool	LGR2	31720.7557287983	174283	0.161140942788191	0.885354913106868	0	5.49428902293693	1.0145651333591	0.98564396421659	0.898250225486406
LGPool	LGR3	29522.2937498353	163796	0.149972790332969	0.832081117190159	0	5.54821388161664	1.07204474409384	0.932796886985588	0.892028188343445
LGPool	LGPool	32665.2289764954	196851	0.165938852108932	1	0	6.02631624415203	1	1	1

obsNumPeaksOverlapTable

	LGR1	LGR2	LGR3	LGPool
LGR1	110704	102960	101496	105998
LGR2	79055	194025	159841	178511
LGR3	80497	165512	183622	173809
LGPool	79185	174283	163796	196851

expNumPeaksTable

	LGR1	LGR2	LGR3	LGPool
LGR1	7458.25065441778	15252.3089482641	14156.8444797592	15743.642145067
LGR2	15369.5476439576	30788.6391853734	28647.4271536346	31712.4858985705
LGR3	14251.9850004457	28619.9603238136	26621.7658668381	29486.2988119775
LGPool	15868.7946415307	31720.7557287983	29522.2937498353	32665.2289764954

obsProportionTable

	LGR1	LGR2	LGR3	LGPool
LGR1	1	0.930047694753577	0.916823240352652	0.95749024425495
LGR2	0.407447493879655	1	0.823816518489885	0.920041231800026
LGR3	0.438384289464225	0.90137347376676	1	0.946558691224363
LGPool	0.402258561043632	0.885354913106868	0.832081117190159	1

expProportionTable

	LGR1	LGR2	LGR3	LGPool
LGR1	0.0673711036134	0.137775590297225	0.127880153199154	0.142213850855136
LGR2	0.079214264367776	0.158683876744612	0.147648123456434	0.163445359611238
LGR3	0.077615890255229	0.155863460390441	0.144981352271722	0.160581514262874
LGPool	0.080613228490232	0.161140942788191	0.149972790332969	0.165938852108932

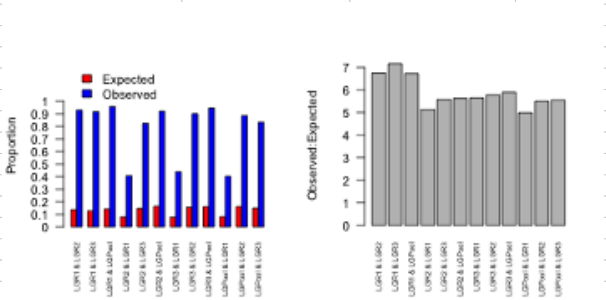
pValueTable

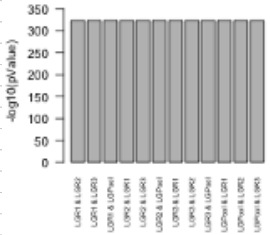
	LGR1	LGR2	LGR3	LGPool
LGR1	0	0	0	0
LGR2	0	0	0	0
LGR3	0	0	0	0
LGPool	0	0	0	0

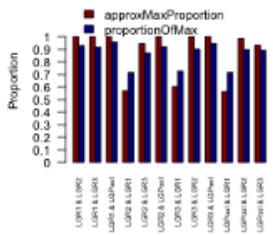
setSizeRatioTable (|A|/|B|)

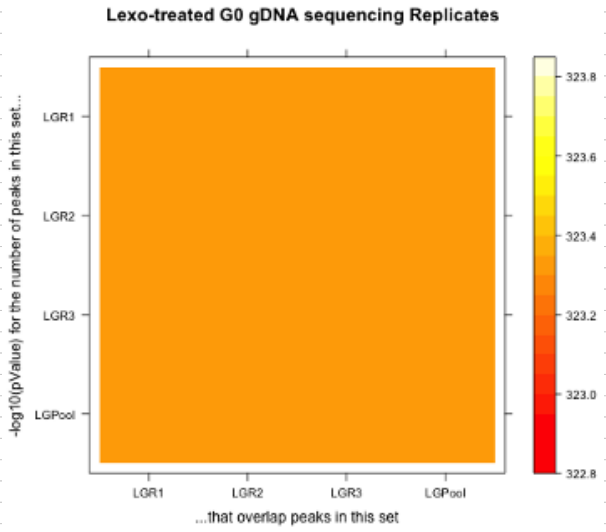
	LGR1	LGR2	LGR3	LGPool
LGR1	1	0.570565648756604	0.602890721155417	0.562374587886269
LGR2	1.75264669749964	1	1.05665443138622	0.98564396421659
LGR3	1.65867538661656	0.94638319804149	1	0.932796886985588
LGPool	1.77817423038011	1.0145651333591	1.07204474409384	1

Overlap_analysis_of_replicates_q1e-3sets_noChrYchrM - LexoG0









Overlap_analysis_of_replicates_q1e-3sets_noChrYchrM - NS-gDNA

Overlap of q<0.001 NS-seq gDNA-controlled peak sets

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportion	fractionOfMax
NSgR1	NSgR1	4809.09134904265	100594	0.047806940265251	1	0	20.9174650051148	1	1	1
NSgR1	NSgR2	6394.38300437596	41533	0.063566246539316	0.412877507604827	0	6.49523182636653	1.05854993160055	0.944688550012923	0.437051457434494
NSgR1	NSgR3	6939.02347904284	25147	0.068980490675814	0.249985088573871	0	3.62399696094828	1.15608012595819	0.864991947829891	0.28900279268615
NSgR1	NSgPool	10873.963524767	79622	0.108097535884516	0.791518380817941	0	7.32226108894421	0.62057520759047	1	0.791518380817941
NSgR2	NSgR1	6391.13049584774	35764	0.06725381980267	0.376344312322424	0	5.59588010653757	0.944688550012923	1	0.376344312322424
NSgR2	NSgR2	7785.63019161614	95030	0.081928129975967	1	0	12.2058199093931	1	1	1
NSgR2	NSgR3	8152.41766969574	65670	0.085787831944604	0.691044933178996	0	8.0552791405805	1.0921356578902	0.915637167210355	0.754714812729132
NSgR2	NSgPool	13248.9494148106	92302	0.139418598493219	0.97129327580764	0	6.96674106830075	0.58625029303261	1	0.97129327580764
NSgR3	NSgR1	6930.41463965376	23076	0.079648036956015	0.265201751462425	0	3.32967090712957	0.86499194782989	1	0.265201751462425
NSgR3	NSgR2	8146.44714080091	42878	0.093623333763931	0.492776941376576	0	5.26339878709193	0.915637167210355	1	0.492776941376576
NSgR3	NSgR3	8395.75442851733	87013	0.096488506642885	1	0	10.3639286666661	1	1	1
NSgR3	NSgPool	13867.0882797557	78512	0.15936800569749	0.902301954880305	0	5.6617509325745	0.536792557588619	1	0.902301954880305
NSgPool	NSgR1	10935.4804852421	70852	0.067462155518526	0.437093610038372	0	6.47909345141421	1.6114082350836	0.62057520759047	0.704336242718254
NSgPool	NSgR2	13330.682978116	70182	0.08223841736552	0.43296030796185	0	5.26469649868747	1.70575607702831	0.58625029303261	0.738524676417973
NSgPool	NSgR3	13962.8610613101	79603	0.086138391968501	0.491079470443806	0	5.70105221633788	1.86291703538552	0.536792557588619	0.914840311217864
NSgPool	NSgPool	22684.9395281827	162098	0.139945832324783	1	0	7.14562186946175	1	1	1

obsNumPeaksOverlapTable

	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	100594	41533	25147	79622
NSgR2	35764	95030	65670	92302
NSgR3	23076	42878	87013	78512
NSgPool	70852	70182	79603	162098

expNumPeaksTable

	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	4809.09134904265	6394.38300437596	6939.02347904284	10873.963524767
NSgR2	6391.13049584774	7785.63019161614	8152.41766969574	13248.9494148106
NSgR3	6930.41463965376	8146.44714080091	8395.75442851733	13867.0882797557
NSgPool	10935.4804852421	13330.682978116	13962.8610613101	22684.9395281827

obsProportionTable

	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	1	0.41287750760482	0.24998508857387	0.791518380817941
NSgR2	0.37634431232242	1	0.69104493317899	0.97129327580764
NSgR3	0.26520175146242	0.49277694137657	1	0.902301954880305
NSgPool	0.43709361003837	0.43296030796185	0.49107947044380	1

expProportionTable

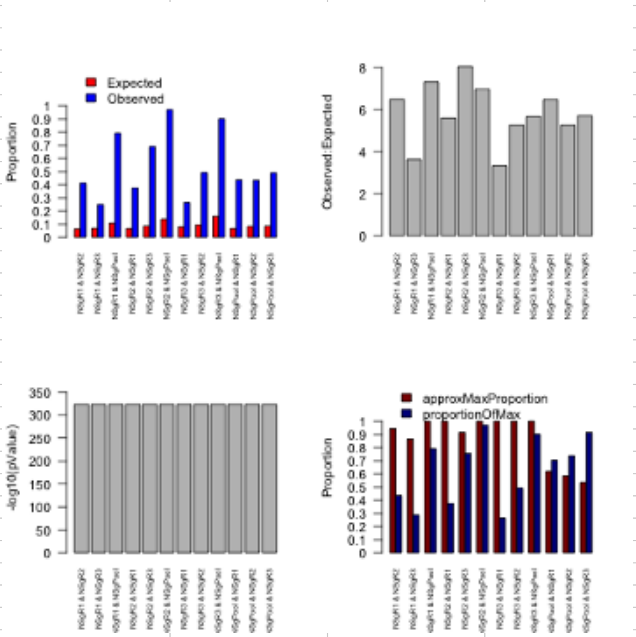
	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	0.04780694026525	0.06356624653931	0.06898049067581	0.108097535884516
NSgR2	0.06725381980267	0.08192812997596	0.08578783194460	0.139418598493219
NSgR3	0.07964803695601	0.09362333376393	0.09648850664288	0.15936800569749
NSgPool	0.06746215551852	0.08223841736552	0.08613839196850	0.139945832324783

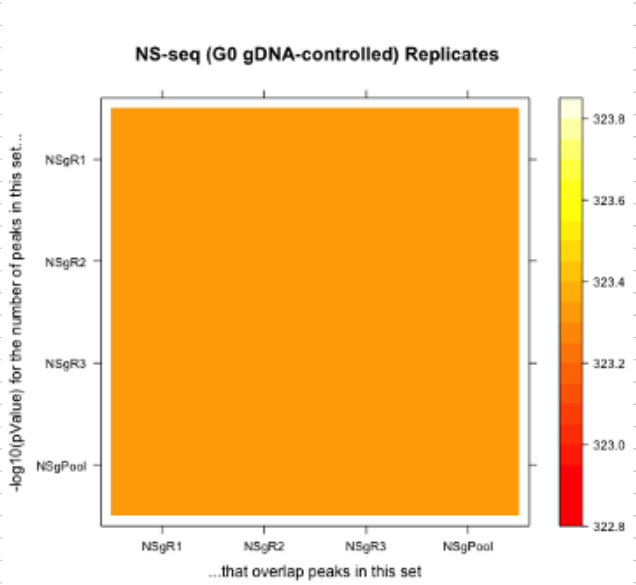
pValueTable

	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	0	0	0	0
NSgR2	0	0	0	0
NSgR3	0	0	0	0
NSgPool	0	0	0	0

setSizeRatioTable

	NSgR1	NSgR2	NSgR3	NSgPool
NSgR1	1	1.05854993160055	1.15608012595819	0.62057520759047
NSgR2	0.94468855001292	1	1.0921356578902	0.58625029303261
NSgR3	0.86499194782989	0.91563716721035	1	0.536792557588619
NSgPool	1.6114082350836	1.70575607702831	1.86291703538552	1


Expected
Observed
Proportion
Observed:Expected
-log10(pValue)
approxMaxProportion
proportionOfMax
Proportion


NS-seq (G0 gDNA-controlled) Replicates
-log10(pValue) for the number of peaks in this set...
NSgR1
NSgR2
NSgR3
NSgPool
...that overlap peaks in this set
323.8
323.6
323.4
323.2
323.0
322.8

Overlap_analysis_of_replicates_q1e-3sets_noChrYchrM - NS-lexo

Overlap of q<0.001 NS-seq lexo-controlled peak sets										
fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
NSIR1	NSIR1	8.06906096187519	4448	0.00181408744646	1	0	551.24134282985	1	1	1
NSIR1	NSIR2	141.825741087526	518	0.03188528351787	0.11645683453237	0	3.65236942199599	0.08728928312106	1	0.11645683453237
NSIR1	NSIR3	235.703354675164	1953	0.05299086211222	0.43907374100719	0	8.28583879381579	0.06568804087781	1	0.43907374100719
NSIR1	NSIPool	228.981308773713	2801	0.05147961078545	0.62972122302158	0	12.2324394729006	0.06655593960886	1	0.62972122302158
NSIR2	NSIR1	142.426331577567	458	0.00279502976190	0.00898797024942	0	3.21569751131706	11.4561600719424	0.08728928312106	0.10296762589928
NSIR2	NSIR2	2199.82045930105	50957	0.04317013284339	1	0	23.1641631409277	1	1	1
NSIR2	NSIR3	3466.68936671562	50450	0.06803166133633	0.99005043468022	0	14.5527893223952	0.75253271110848	1	0.99005043468022
NSIR2	NSIPool	3379.50912797238	50654	0.06632080240148	0.99405381007516	0	14.9885672983494	0.76247549789768	1	0.99405381007516
NSIR3	NSIR1	237.063187931825	1921	0.00350094792704	0.028369318014	0	8.10332475809127	15.2234712230216	0.06568804087781	0.43187949640287
NSIR3	NSIR2	3471.98674169038	31397	0.05127428215273	0.46367073278790	0	9.04294927828959	1.32884588967168	0.75253271110848	0.61614694742626
NSIR3	NSIR3	5337.02274229411	67714	0.07881712411457	1	0	12.6875981740511	1	1	1
NSIR3	NSIPool	5211.56484486646	57787	0.07696436253753	0.85339811560386	0	11.0882243088507	1.01321243135671	0.98695986059012	0.86467357962622
NSIPool	NSIR1	230.28381809058	2711	0.00344576346441	0.04056500725711	0	11.7724294415409	15.0249550359712	0.06655593960886	0.60948741007194
NSIPool	NSIR2	3384.40076866104	31623	0.05064118101870	0.47317861471472	0	9.34375157127474	1.31151755401613	0.76247549789768	0.62058205938340
NSIPool	NSIR3	5211.14523721539	59424	0.07797497025654	0.88916820038604	0	11.4032515493185	0.98695986059012	1	0.88916820038604
NSIPool	NSIPool	5088.0616922763	66831	0.07613325690587	1	0	13.1348643239625	1	1	1

obsNumPeaksOver

	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	4448	518	1953	2801
NSIR2	458	50957	50450	50654
NSIR3	1921	31397	67714	57787
NSIPool	2711	31623	59424	66831

expNumPeaksTable

	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	8.06906096187519	141.825741087526	235.703354675164	228.981308773713
NSIR2	142.426331577567	2199.82045930105	3466.68936671562	3379.50912797238
NSIR3	237.063187931825	3471.98674169038	5337.02274229411	5211.56484486646
NSIPool	230.28381809058	3384.40076866104	5211.14523721539	5088.0616922763

obsProportionTable

	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	1	0.11645683453237	0.43907374100719	0.62972122302158
NSIR2	0.00898797024942	1	0.99005043468022	0.99405381007516
NSIR3	0.028369318014	0.46367073278790	1	0.85339811560386
NSIPool	0.04056500725711	0.47317861471472	0.88916820038604	1

expProportionTable

	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	0.00181408744646	0.03188528351787	0.05299086211222	0.05147961078545
NSIR2	0.00279502976190	0.04317013284339	0.06803166133633	0.06632080240148
NSIR3	0.00350094792704	0.05127428215273	0.07881712411457	0.07696436253753
NSIPool	0.00344576346441	0.05064118101870	0.07797497025654	0.07613325690587

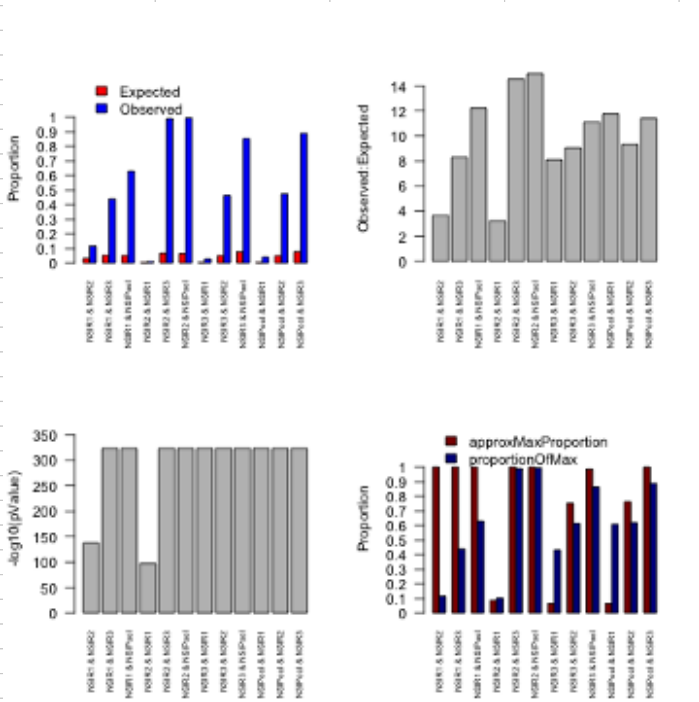
pValueTable

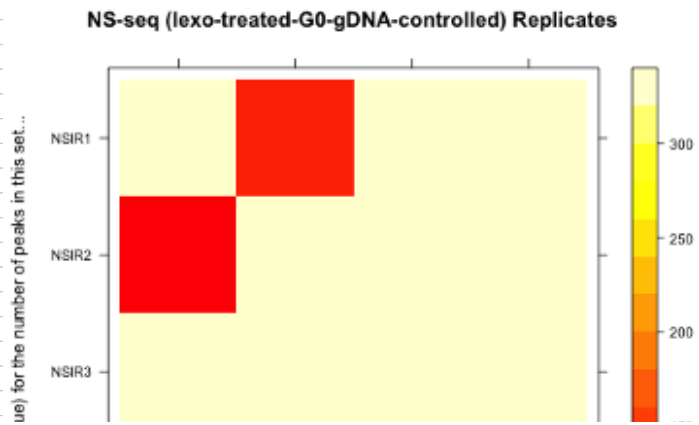
	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	0	0	0	0
NSIR2	0	0	0	0
NSIR3	0	0	0	0
NSIPool	0	0	0	0

setSizeRatioTable

Overlap_analysis_of_replicates_q1e-3sets_noChrYchrM - NS-lexo

	NSIR1	NSIR2	NSIR3	NSIPool
NSIR1	1	0.08728928312106	0.06568804087781	0.06655593960886
NSIR2	11.4561600719424	1	0.75253271110848	0.76247549789768
NSIR3	15.2234712230216	1.32884588967168	1	1.01321243135671
NSIPool	15.0249550359712	1.31151755401613	0.98695986059012	1





All sets were q<0.001, no chrY, no chrM							
Set	**numPeaks**						
Lexo G0 Rep1	110704						
Lexo G0 Rep2	194025						
Lexo G0 Rep3	183622						
Lexo G0 Pool	196851						
NS gDNA Rep1	100594						
NS gDNA Rep2	95030						
NS gDNA Rep3	87013						
NS gDNA Pool	162098						
NS lexo Rep1	4448						
NS lexo Rep2	50957						
NS lexo Rep3	67714						
NS lexo Pool	66831						

OverlapAnalyses - Notes/Tables

For any analysis with Cadoret or Karnani (in either direction):
p-values here should not be used or considered. The p-value is calculated here assuming genome-wide - not 1% ENCODE.
I will have to isolate the features/peaks in other sets for only the ENCODE regions as well as get precise boundaries for the ENCODE regions
to do those analyses correctly

For numPeaks, Proportions, and pValues tables, read as:
this is the numPeaks/proportionPeaks/p-value in/for the ROW set that overlapped peaks in the COLUMN set

obsNumPeaksOverlapTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	196851	62230	12513	215	39	63	26	36741	52566	83482	7645	72554	23865	14254	3097	2443	4674	29877
NS-gDNA	76265	162098	60434	152	34	24	10	15751	25233	60768	3814	56600	13405	40759	2483	1981	2635	15334
NS-lexo	13492	62357	66831	6	21	3	3	144	119	13871	137	23718	39	17477	1250	944	633	3449
Cadoret	204	119	6	282	18	16	6	57	70	131	32	207	99	125	8	6	7	50
Karnani-BrIP	38	34	21	18	815	104	104	16	8	231	3	67	2	549	47	40	26	64
Karnani-Lexo	54	21	3	16	107	320	150	13	12	86	4	50	13	170	20	15	10	33
Karnani-Ori	23	10	3	6	107	150	150	3	4	38	2	16	4	75	11	9	7	16
Aladjem-K562	46606	17866	155	63	16	14	3	62971	29120	29873	1344	13014	3204	695	260	201	230	1771
Aladjem-MCF7	72387	29179	121	85	8	13	4	28025	94195	46605	2143	24923	6095	1747	424	329	543	3932
Bubble	47019	33098	11385	132	213	85	38	17461	22554	123274	6172	50306	11797	106901	25660	23159	12866	36338
ORC	7650	3389	122	30	3	4	2	1273	1958	6479	13603	4632	4169	2803	269	224	143	1764
G4	174050	93270	24383	455	77	61	21	17971	37913	138897	10075	358917	35441	0	110	78	635	555
CpG	26189	12600	33	99	2	13	4	2972	5594	14624	4188	14937	28510	508	213	179	108	6778
ATtract	15985	49947	20614	176	1034	304	142	715	1936	614008	3637	0	621	2550407	54882	64689	57558	4556
IR	3222	2732	1396	9	52	25	15	264	429	35800	298	106	222	62347	145766	120921	10459	2567
Cruciform	2566	2261	1112	11	56	28	21	204	332	40922	278	77	188	91664	134421	169106	30090	3055
Bends	5093	2818	661	7	26	10	7	234	546	14112	146	584	113	27540	6664	9271	54711	2334
Z-DNA	53391	26886	3713	77	77	44	23	2108	4656	60939	2419	551	10462	4429	2607	2531	3591	187348

expNumPeaksTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7
LexoG0	32665.2289764954	27267.5892545402	13128.2891086475	53.1615069100934	162.431578483439	57.6077953858564	27.6924362430985	6764.97925600409	10784.0214041714
NS-gDNA	27180.4241452094	22684.9395281827	10900.9557712427	44.1634438278405	134.85138347911	47.8841083861532	23.0110440417794	5670.58763329389	9027.88947202306
NS-lexo	12972.6449440302	10806.2619557496	5088.0616922763	20.7199234797608	62.8312308114726	22.6003761674272	10.8249958267836	2913.83035114607	4581.61748609277
Cadoret	52.2143990055918	43.5157473379895	20.5949502967825	0.08375826140614	0.25443662341978	0.09122147659747	0.04372920682537	11.5189394387938	18.1657529601023
Karnani-BrIP	159.545458197888	132.880086167456	62.4552783749927	0.25444891520499	0.77112531432778	0.27768604507884	0.13296635132131	36.0543775644132	56.6347060816736
Karnani-Lexo	56.5816687404218	47.1820075131979	22.4641383143753	0.09122179077081	0.27767358709108	0.09917551042517	0.04758825725857	12.2174175323516	19.3365901509693
Karnani-Ori	27.1987515219181	22.6732952011135	10.7595756375308	0.04372868367386	0.13295833739487	0.0475875240447	0.02282214555343	5.94323390515559	9.38759981347283
Aladjem-K562	6682.42415909547	5619.350862807	2912.8051400783	11.5847606024175	36.258646713148	12.2871875997001	5.97726605516514	973.439792505815	1666.13725059251
Aladjem-MCF7	10682.8251106954	8971.85249196557	4593.07783754157	18.3217005418614	57.1181388084246	19.502521853396	9.468302836886	1670.89277964523	2814.45492555645
Bubble	62804.785051379	51946.5989499316	22590.1691126769	93.9315085153806	276.936817081224	104.898754791455	49.5997185556814	17799.0341833	27038.0512042924
ORC	1904.41748977103	1593.2935383493	785.069919524683	3.16082978744761	9.73240907224702	3.40210464940111	1.64154202455618	358.84077078564	581.936284172455
G4	30937.9954024469	26159.497479058	14276.5432996783	56.1050882848822	178.419695459124	58.6356617845314	28.7603843503231	3077.16010628405	5833.20315769293
CpG	3840.45999627903	3215.10533500397	1594.54842322259	6.40966636977646	19.7781261750107	6.88584993543077	3.32597280316565	703.086910909465	1146.49904899919
ATtract	274713.760661426	232325.476542218	127002.356401133	498.915279308164	1587.39637059917	521.172165844461	255.699102907635	26900.9212775292	51239.8478751253
IR	12320.1526936967	10417.2253892074	5685.04988948681	22.3416925567884	71.0482116856621	23.3495712368899	11.4527541107104	1225.68974143543	2323.2981952393
Cruciform	14274.6143735832	12070.8888439437	6592.89085920838	25.9046056991699	82.3988677193914	27.0669298008214	13.2778146485353	1409.34550095311	2677.67516136821
Bends	5429.18126420099	4571.98088614263	2402.41751987098	9.52421247983302	29.9373237800502	10.06216731306	4.90559646541392	726.010243157576	1268.33451504289
Z-DNA	15802.8136654015	13364.0112822639	7303.38667179993	28.6924952792375	91.2826802026386	29.9749856923903	14.7057313398379	1551.76746105421	2953.21261036099

	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	63231.1966007177	1936.62043208056	30475.3180591127	3900.11780505419	196851	12378.2296218771	14311.1606608134	5500.3918899638	15816.6713061363
NS-gDNA	52132.1062619035	1615.05616246777	25685.9101727156	3254.61162883204	162098	10432.8746946614	12063.1078159218	4617.14149115921	13332.9727017437
NS-lexo	22473.9054180062	788.880253424829	13896.3115682622	1600.12014080708	66831	5644.13150841947	6531.40716929817	2405.07252677776	7223.12888932459
Cadoret	92.8844394938533	3.15701364051645	54.2814358485302	6.39326792417929	282	22.0471022488104	25.5082379393813	9.4772288483973	28.2060317023543
Karnani-BrIP	273.862989272727	9.72112850048434	172.628640714198	19.7284789884283	815	70.1147800902968	81.1419979946637	29.7910799208943	89.7393736917831
Karnani-Lexo	103.729789434803	3.39800890838923	56.7299503980961	6.86825687982156	320	23.0417707249493	26.6528690517636	10.0125643952624	29.4668797665451
Karnani-Ori	49.0462360699328	1.63954053604306	27.8252174453187	3.31742396935516	150	11.3016066038752	13.0744936774408	4.88133840372691	14.4562317785214
Aladjem-K562	17701.1985469664	360.455534690227	2994.15127744983	705.295417030837	20801.6101463178	1216.43964244599	1395.7110449043	726.556861894228	1534.17492552097
Aladjem-MCF7	26966.1800896447	586.223416580415	5692.04802836437	1153.38302630322	39735.2032692531	2312.34577662759	2659.33922845387	1272.91229044277	2928.06532633803
Bubble	65936.7405802985	4140.55688698441	96421.7092816364	8579.27662600849	123274	39160.3429098322	45400.1632556035	15229.6780239916	50273.8125344016
ORC	4099.35070650618	109.81186577491	1496.0087077143	219.361391919259	10528.3890175402	607.663128279579	701.627657136453	286.137844958637	774.713205532548
G4	98550.1788350772	1544.40213501962	2577.22431451379	2942.9736869375	15534.7185571998	1049.21762447152	1126.39682752648	1975.30935296281	1177.05028122109
CpG	8505.40226710468	219.658524390089	2854.61784115009	437.730825467242	20070.3701157653	1159.53372620806	1338.2002105035	557.065454145723	1477.09686344872
ATtract	879983.364009645	13676.7227073403	19547.7706956818	26036.8014356894	114058.314023492	7961.56306654898	8424.29040711226	17129.2126475065	8699.50099056614
IR	39241.2387641607	615.039188916379	1028.67742501115	1172.02107276964	6203.23344546811	418.783949336473	449.676620945252	786.901218807041	469.972035740251
Cruciform	45591.6813891187	711.669863660878	1106.71812446268	1355.51741301328	6577.8669840715	450.64263809673	480.756000498783	901.258369541999	499.816338421485
Bends	15134.5703869119	287.209207623257	1920.57780219121	558.394864178782	13235.4996057116	780.375041411563	891.867815606285	527.2959642578 7	977.575073470258
Z-DNA	50570.7993279259	787.123061856795	1158.43166866555	1498.72787678854	6804.18168572545	471.773770555525	500.656951957899	989.52948738137	518.264676599045

obsProportionTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7
LexoG0	1	0.31612742632752	0.0635658442172	0.00109219663603	0.00019811938979	0.00032003901428	0.00013207959319	0.18664370513738	0.26703445753387
NS-gDNA	0.47048698935212	1	0.37282384730224	0.00093770435168	0.00020974965761	0.00014805858184	0.00006169107576	0.09716961344371	0.15566509148786
NS-lexo	0.20188235998264	0.93305501937723	1	0.00008977869551	0.00031422543430	0.00004488934775	0.00004488934775	0.00215468869237	0.00178061079439
Cadoret	0.72340425531914	0.42198581560283	0.02127659574468	1	0.06382978723404	0.05673758865248	0.02127659574468	0.20212765957446	0.24822695035461
Karnani-BrIP	0.04662576687116	0.04171779141104	0.02576687116564	0.02208588957055	1	0.12760736196319	0.12760736196319	0.01963190184049	0.00981595092024
Karnani-Lexo	0.16875	0.065625	0.009375	0.05	0.334375	1	0.46875	0.040625	0.0375
Karnani-Ori	0.15333333333333	0.06666666666666	0.02	0.04	0.71333333333333	1	1	0.02	0.02666666666666
Aladjem-K562	0.74011846723094	0.28371790189134	0.00246145050896	0.00100046052945	0.00025408521382	0.0002223245621	0.00004764097759	1	0.46243508916803
Aladjem-MCF7	0.76848028026965	0.30977228090663	0.00128456924465	0.00090238335368	0.00008493019799	0.00013801157173	0.00004246509899	0.29752109984606	1
Bubble	0.38141862842124	0.26849132826062	0.09235524117007	0.00107078540487	0.00172785826695	0.00068952090465	0.00030825640443	0.14164381783668	0.18295828804127
ORC	0.56237594648239	0.24913621995148	0.00896860986547	0.00220539586855	0.00022053958685	0.00029405278247	0.00014702639123	0.09358229802249	0.14393883702124
G4	0.48493105648381	0.25986509415825	0.06793492645932	0.00126770256075	0.00021453427951	0.00016995572792	0.00005850934895	0.05007007191077	0.10563166414519
CpG	0.91858996843212	0.44195019291476	0.00115748860049	0.00347246580147	0.00007015082427	0.00045598035776	0.00014030164854	0.10424412486846	0.19621185548930
ATtract	0.00626762708853	0.01958393307421	0.00808263151724	0.00006900859353	0.00040542548699	0.00011919666155	0.00005567738796	0.00028034741121	0.00075909452883
IR	0.02210391998134	0.01874236790472	0.00957699326317	0.00006174279324	0.00035673613874	0.00017150775901	0.00010290465540	0.00181112193515	0.00294307314462
Cruciform	0.01517391458611	0.01337031211193	0.00657575721736	0.00006504795808	0.00033115324116	0.00016557662058	0.00012418246543	0.00120634394994	0.00196326564403
Bends	0.09308914112335	0.05150700955932	0.01208166547860	0.00012794502019	0.00047522436073	0.00018277860028	0.00012794502019	0.00427701924658	0.00997971157536
Z-DNA	0.28498302623993	0.14350833742554	0.01981873305292	0.00041099985054	0.00041099985054	0.00023485705745	0.00012276618912	0.01125178811623	0.02485214680701

	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	0.42408725381125	0.03883648038364	0.36857318479459	0.12123382659981	0.07241009697690	0.01573271154324	0.01241040177596	0.02374384686895	0.15177469253394
NS-gDNA	0.37488432923293	0.02352897629828	0.34917148885242	0.08269688706831	0.25144665572678	0.01531789411343	0.01222100210983	0.01625559846512	0.09459709558415
NS-lexo	0.20755338091604	0.00204994688093	0.35489518337298	0.00058356152085	0.26151037692089	0.01870389489907	0.01412518142778	0.00947165237689	0.05160778680552
Cadoret	0.46453900709219	0.11347517730496	0.73404255319148	0.35106382978723	0.44326241134751	0.02836879432624	0.02127659574468	0.02482269503546	0.17730496453900
Karnani-BrIP	0.28343558282208	0.00368098159509	0.08220858895705	0.00245398773006	0.67361963190184	0.05766871165644	0.04907975460122	0.03190184049079	0.07852760736196
Karnani-Lexo	0.26875	0.0125	0.15625	0.040625	0.53125	0.0625	0.046875	0.03125	0.103125
Karnani-Ori	0.25333333333333	0.01333333333333	0.10666666666666	0.02666666666666	0.5	0.07333333333333	0.06	0.04666666666666	0.10666666666666
Aladjem-K562	0.47439297454383	0.02134315796160	0.20666656079782	0.05088056406917	0.01103682647567	0.00412888472471	0.00319194549872	0.00365247494878	0.02812405710565
Aladjem-MCF7	0.49477148468602	0.02275067678751	0.26458941557407	0.06470619459631	0.01854663198683	0.00450130049365	0.00349275439248	0.0057646371888	0.04174319231381
Bubble	1	0.05006732968833	0.40808280740464	0.09569738955497	0.86718204974284	0.20815419309830	0.18786605448026	0.10436912893229	0.29477424274380
ORC	0.47629199441299	1	0.34051312210541	0.306476512534	0.20605748731897	0.01977504962140	0.01646695581856	0.01051238697346	0.12967727707123
G4	0.3869891924874	0.02807055670252	1	0.09874427792498	0	0.00030647754216	0.00021732043898	0.00176921126611	0.00154631850817
CpG	0.51294282707821	0.14689582602595	0.52392143107681	1	0.01781830936513	0.00747106278498	0.00627849877236	0.00378814451069	0.23774114345843
ATtract	0.24074902554768	0.00142604690153	0	0.00024349054876	1	0.02151891835303	0.02536418697094	0.02256816265011	0.00178638154616
IR	0.24559911090377	0.00204437248741	0.00072719289820	0.00152298890001	0.42771977004239	1	0.82955558909485	0.07175198605984	0.01761041669525
Cruciform	0.24199023097938	0.00164393930434	0.00045533570659	0.00111172873818	0.54205054817688	0.79489196125507	1	0.17793573261741	0.01806559199555
Bends	0.25793716071722	0.00266856756411	0.01067427025643	0.00206539818318	0.50337226517519	0.12180365922757	0.16945404032095	1	0.04266052530569
Z-DNA	0.32527168691419	0.01291180049960	0.00294105087857	0.05584260307022	0.02364049789696	0.01391528065418	0.01350961846403	0.01916753848453	1

expProportionTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7
LexoG0	0.16593885210893	0.13851892677477	0.06669150326209	0.00027005962332	0.00082514987723	0.00029264669920	0.00014067714282	0.03436598877325	0.05478266000259
NS-gDNA	0.16767896053751	0.13994583232478	0.06724916884380	0.00027244903594	0.00083191269157	0.00029540221585	0.00014195760615	0.03498246513401	0.05569402134525
NS-lexo	0.19411119007691	0.16169535029776	0.07613325690587	0.00031003461686	0.00094015098998	0.00033817204841	0.00016197566738	0.04359998131325	0.06855527354211
Cadoret	0.18515744328224	0.15431116077301	0.07303173864107	0.00029701511136	0.00090225752985	0.00032348041346	0.00015506810931	0.04084730297444	0.06441756368830
Karnani-BrIP	0.19576129840231	0.16304305051221	0.07663224340489	0.00031220725792	0.00094616602985	0.00034071907371	0.00016314889732	0.04423850007903	0.06949043691002
Karnani-Lexo	0.17681771481381	0.14744377347874	0.07020043223242	0.00028506809615	0.00086772995966	0.00030992347007	0.00014871330393	0.03817942978859	0.06042684422177
Karnani-Ori	0.18132501014612	0.15115530134075	0.07173050425020	0.00029152455782	0.00088638891596	0.00031725016029	0.00015214763702	0.03962155936770	0.06258399875648
Aladjem-K562	0.10611907320981	0.08923712284713	0.04625629480361	0.00018396977342	0.00057579912520	0.00019512454303	0.00009492093273	0.01545854111425	0.02645880247403
Aladjem-MCF7	0.11341180647269	0.09524765106391	0.04876137626775	0.00019450820682	0.00060638185475	0.00020704413029	0.00010051810432	0.01773865682515	0.02987902675892
Bubble	0.50947308476547	0.42139136354731	0.18325169226825	0.00076197339678	0.00224651440759	0.00085093981530	0.00040235344481	0.14438595472930	0.21933295913406
ORC	0.13999981546504	0.11712809956254	0.05771299856830	0.00023236269848	0.00071546049196	0.00025009958460	0.00012067499996	0.02637953177869	0.04277999589593
G4	0.08619818900315	0.07288453174148	0.03977672637316	0.00015631772327	0.00049710572488	0.00016336830460	0.00008013101733	0.00857345878374	0.01625223424271
CpG	0.13470571716166	0.11277114468621	0.05592944311548	0.00022482168957	0.00069372592686	0.00024152402439	0.00011665986682	0.02466106316764	0.04021392665728
ATtract	0.10771369458342	0.09109349078096	0.04979689767207	0.00019562182793	0.00062240903926	0.00020434862586	0.00010025815601	0.01054769739791	0.02009085133279
IR	0.08452007116677	0.07146539926462	0.03900120665646	0.00015327094491	0.00048741278271	0.00016018530546	0.00007856944768	0.00840861203185	0.01593854667919
Cruciform	0.08441222885990	0.07138060650682	0.03898673529743	0.00015318560961	0.00048726164488	0.00016005895592	0.00007851770279	0.00833409518853	0.01583430015119
Bends	0.09923381521450	0.08356602668828	0.04391105115737	0.00017408222258	0.00054719021366	0.00018391488572	0.00008966380554	0.01326991360343	0.02318244073482
Z-DNA	0.08435005265816	0.07133255376232	0.03898299779981	0.00015315079573	0.00048723594702	0.00015999629402	0.00007849419977	0.00828280772174	0.01576324599334

	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	0.32121348939409	0.00983800149392	0.15481413891274	0.01981253742706	1	0.062881212805	0.07270047223947	0.02794190473995	0.08034844276196
NS-gDNA	0.32160857173995	0.00996345520899	0.15845914306601	0.02007804925928	1	0.06436152632766	0.07441860982814	0.02848364255672	0.0822525429169
NS-lexo	0.33627965192809	0.01180410667841	0.20793212084604	0.02394278315163	1	0.08445379402402	0.09773020259008	0.03598737901240	0.10808051487071
Cadoret	0.32937744501366	0.01119508383161	0.19248736116500	0.0226711628517	1	0.07818121364826	0.09045474446589	0.03360719449786	0.10002138901544
Karnani-BrIP	0.33602820769659	0.01192776503126	0.21181428308490	0.02420672268518	1	0.08603040501876	0.09956073373578	0.03655347229557	0.11010966097151
Karnani-Lexo	0.32415559198375	0.01061877783871	0.17728109499405	0.02146330274944	1	0.07200553351546	0.08329021578676	0.03128926373519	0.09208399927045
Karnani-Ori	0.32697490713288	0.01093027024028	0.18550144963545	0.02211615979570	1	0.07534404402583	0.08716329118293	0.03254225602484	0.09637487852347
Aladjem-K562	0.28110080111426	0.00572415135046	0.04754809797287	0.01120032105303	0.33033634762537	0.01931745791627	0.02216434620546	0.01153795972581	0.02436319775009
Aladjem-MCF7	0.28628037676781	0.00622350885482	0.06042834575470	0.01224463109828	0.42183983512132	0.02454849807980	0.02823227590056	0.01351358660696	0.03108514598798
Bubble	0.53487954134934	0.03358824153499	0.78217393190483	0.06959518329906	1	0.31766911846644	0.36828660752148	0.12354331021944	0.40782170234114
ORC	0.30135637039669	0.00807262116995	0.10997638077735	0.01612595691533	0.77397552139529	0.04467125841943	0.05157889121050	0.02103490737033	0.05695164342663
G4	0.27457651444505	0.00430295064045	0.00718055794101	0.00819959402017	0.04328220328711	0.00292328762491	0.00313832119271	0.00550352686822	0.00327944979262
CpG	0.29833048990195	0.00770461327218	0.10012689726938	0.01535358910793	0.70397650353438	0.04067112333244	0.04693792390401	0.01953930039094	0.05180978125039
ATtract	0.34503644477514	0.0053625647621 5	0.00766456910433	0.01020888094946	0.04472161267730	0.00312168334957	0.00330311609367	0.00671626632435	0.00341102458963
IR	0.26920707685029	0.0042193597198	0.00705704639635	0.00804042830817	0.04255610667417	0.00287298786641	0.00308492118151	0.00539838658402	0.00322415402590
Cruciform	0.26960416182228	0.00420842467837	0.00654452310658	0.00801578544234	0.03889789235196	0.00266485303949	0.00284292692452	0.00532954696783	0.00295563929382
Bends	0.27662755911812	0.00524956969573	0.03510405224161	0.01020626316789	0.24191660919580	0.01426358577638	0.01630143509726	0.00963784182811	0.01786798035989
Z-DNA	0.26992975280187	0.00420139559459	0.00618331484011	0.00799970043335	0.03631841111581	0.00251816817129	0.00267233678479	0.00528177235615	0.00276632083928

pValueTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	0	0	0.99999998889259	0	1	0.21612130805891	0.57778502247641	0	0	0	0	0	0	1	1	1	1	0

NS-gDNA	0	0	0	0	1	0.99989486582404	0.99803645262654	0	0	0	0	0	0	1	1	1	1	0
NS-lexo	0.00000020965063	0	0	0.9998486229073	0.99999999914462	0.99999966385382	0.99439542983381	1	1	1	1	0	1	1	1	1	1	1
Cadoret	0	0	0.99988972470178	0	0	0	0.00000000000005	0	0	0.0000092779136	0	0	0	1	0.99961941109732	0.99999802032802	0.73338816051831	0.00002506008660
Karnani-BrIP	1	1	0.99999999964079	0	0	0	0	0.99988579465140	1	0.99926844119917	0.98770088641200	1	0.99999952194310	1	0.99849810247631	0.99999989942580	0.72446159412399	0.99836792980757
Karnani-Lexo	0.61429414795394	0.99999551184755	0.9999997972586	0	0	0	0	0.33994607153123	0.95268803219380	0.98139220062551	0.25484792734925	0.81844106122754	0.01022223586066	1	0.70115218149244	0.99184630002608	0.41854683961788	0.21467112962695
Karnani-Ori	0.78085457782465	0.99873415897237	0.99527434127228	0.00000000000005	0	0	0	0.84907816548504	0.96132914007649	0.96881237242629	0.22638718712149	0.99395208621649	0.23917837857349	1	0.45806231699949	0.85090514499781	0.11777268054219	0.27735944086105
Aladjem-K562	0	0	1	0	0.99986818309851	0.25454340042855	0.84676864626550	0	0	0	0	0	0	1	1	1	1	0.00000000102834
Aladjem-MCF7	0	0	1	0	1	0.91917980147331	0.95893179667686	0	0	0	0	0	0	1	1	1	1	0
Bubble	1	1	1	0.00008390237959	0.99996363498984	0.97389467125740	0.94702961832874	0.99694137198677	1	0	0	1	0	1	1	1	1	1
ORC	0	0	1	0	0.98745992350913	0.25619561850016	0.22742500980487	0	0	0	0	0	0	1	1	1	1	0
G4	0	0	0	0	1	0.34725744767800	0.91698818384550	0	0	0	0	0	0	1	1	1	1	1
CpG	0	0	1	0	0.99999944635117	0.01125989151076	0.24214888736337	0	0	0	0	0	0	1	1	1	1	0
ATtract	1	1	1	1	1	1	0.99999999999999	1	1	1	1	1	1	0	0	0	0	1
IR	1	1	1	0.99878288280470	0.98893769293631	0.31824555189122	0.11873980002590	1	1	1	1	1	1	0	0	0	0	0
Cruciform	1	1	1	0.99917210535176	0.99868331011506	0.38019107917546	0.01738365276956	1	1	1	1	1	1	0	0	0	0	0
Bends	0.99999936194505	1	1	0.73386263145256	0.72901546967403	0.42473821999001	0.12369014221448	1	1	1	1	1	1	0	0	0	0	0
Z-DNA	0	0	1	0.00000000000002	0.92837407300447	0.00617856139915	0.01584949485155	0	0	0	0	0	1	0	1	0	0	0

setSizeRatioTable

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	1	1.21439499561993	2.94550433182206	698.053191489362	241.534969325153	615.159375	1312.34	3.12605802671071	2.0898243006529	1.59685740707692	14.4711460707197	0.54845827865495	6.90462995440196	0.07718415139230	1.35045895476311	1.16406869064374	3.59801502440094	1.05072378674979
NS-gDNA	0.82345530375766	1	2.42549116428005	574.815602836879	198.893251533742	506.55625	1080.65333333333	2.57416906194915	1.72087690429428	1.3149407012022	11.916341983386	0.45163087844822	5.68565415643634	0.06355769883003	1.11204258880672	0.95855853724882	2.96280455484272	0.86522407498345
NS-lexo	0.33950043433866	0.41228762847166	1	236.989361702128	82.001226993865	208.846875	445.54	1.06129805783615	0.70949625776315	0.54213378327952	4.91296037638756	0.18620182381999	2.3441248684672	0.02620405292174	0.45848140169861	0.39520182607358	1.22152766354115	0.35672118197151
Cadoret	0.00143255558772	0.00173968833668	0.00421959868923	1	0.346012269938650	0.88125	1.88	0.00447825189372	0.00299378947927	0.00228758700131	0.02073072116444	0.00078569697172	0.00989126622237	0.00011057058736	0.00193460752164	0.00166759310728	0.00515435652793	0.00150522023186
Karnani-BrIP	0.00414018724822	0.00502782267517	0.01219493947419	2.89007092198582	1	2.546875	5.43333333333333	0.01294246557939	0.00865226392059	0.00661128867401	0.05991325442917	0.00227071997147	0.02858646089091	0.00031955683935	0.00559115294375	0.00481946234905	0.0148964559294	0.00435019322330
Karnani-Lexo	0.00162559499316	0.00197411442460	0.00478819709416	1.13475177304965	0.39263803680981	1	2.13333333333333	0.00508170427657	0.00339720791974	0.00259584340574	0.02352422259795	0.00089157103174	0.01122413188355	0.00012547017005	0.00219529931534	0.00189230423521	0.00584891520900	0.00170805132694
Karnani-Ori	0.00076199765304	0.00092536613653	0.00224446738788	0.53191489361702	0.18404907975460	0.46875	1	0.00238204887964	0.00159244121237	0.00121680159644	0.01102697934279	0.00041792392112	0.00526131182041	0.00005881414221	0.00102904655406	0.00088701761025	0.00274167900422	0.00080064905950
Aladjem-K562	0.31989169473358	0.38847487322483	0.94224237255166	223.301418439716	77.2650306748466	196.784375	419.806666666667	1	0.66851743723127	0.51082142219770	4.62919944129971	0.17544724824959	2.20873377762189	0.02469056899545	0.43200060370731	0.37237590623632	1.1509751238325	0.33611781284027
Aladjem-MCF7	0.47850912619189	0.5810990882059	1.4094507040146	334.024822695035	115.576687116564	294.359375	627.966666666667	1.49584729478649	1	0.76411084251342	6.9245754612953	0.26244229167189	3.30392844615924	0.03693332083859	0.64620693440171	0.55701749198727	1.72168302535139	0.50278092106667
Bubble	0.62622999121162	0.76049056743451	1.84456315183074	437.141843971631	151.256441717791	385.23125	821.826666666667	1.95763129059408	1.30871065343171	1	9.06226567668897	0.34346102302203	4.32388635566468	0.04833503044808	0.84569789937296	0.72897472591155	2.2531849171099	0.65799474774217
ORC	0.06910302716267	0.08391837036854	0.20354326584968	48.2375886524823	16.6907975460123	42.509375	90.6866666666667	0.21602007273189	0.14441318541323	0.11034768077615	1	0.03790012732748	0.47713083128726	0.00533365851019	0.09332080183307	0.08044067034877	0.24863372996289	0.07260819437624
G4	1.82329274425835	2.21419758417747	5.37051667639269	1272.75531914894	440.388957055215	1121.615625	2392.78	5.6997189182322	3.81036148415521	2.91153852393854	26.3851356318459	1	12.5891616976499	0.14072930320533	2.4622820136383	2.12243799746904	6.56023468772276	1.91577705660055
CpG	0.14483035392251	0.17588125701736	0.42659843485807	101.099290780142	34.9815950920245	89.09375	190.066666666667	0.45274809039081	0.30266999309942	0.23127342343073	2.09586120708667	0.07943340660932	1	0.01117860796335	0.19558744837616	0.16859248045604	0.52110178940249	0.15217669790977
ATtract	12.9560276554348	15.7337351478735	38.1620355822897	9043.99645390071	3129.33374233129	7970.021875	17002.7133333333	40.5012942465579	27.0758214342587	20.6889287278745	187.488568698081	7.10584062610576	89.4565766397755	1	17.4965835654405	15.0817061488061	46.6159821608086	13.6132064393535
IR	0.74048899929388	0.89924613505410	2.18111355508671	516.900709219858	178.853987730061	455.51875	971.773333333333	2.31481157993362	1.5474919050905	1.18245534338141	10.7157244725428	0.40612732191565	5.11280252542967	0.05715401502583	1	0.86198005984412	2.66429054486301	0.77804940538463
Cruciform	0.85905583410803	1.0432331058989	2.53035268064222	599.666666666667	207.492024539877	528.45625	1127.37333333333	2.68545838560607	1.79527575773661	1.37178967178805	12.4315224582813	0.47115628404338	5.93146264468607	0.06630549555423	1.16011964381269	1	3.09089579791998	0.90263039904349
Bends	0.27793102397244	0.33751804463966	0.81864703505858	194.010638297872	67.1300613496932	170.971875	364.74	0.86882850836099	0.58082700780296	0.44381621428687	4.02198044548997	0.15243357099273	1.91901087337776	0.02145187023090	0.37533444013007	0.32353080316487	1	0.29202873796357
Z-DNA	0.95172490868728	1.15576996631667	2.80330984124134	664.354609929078	229.874846625767	585.4625	1248.98666666667	2.97514729002239	1.98893784171134	1.51976896993689	13.772550172756	0.52198140517166	6.57130831287268	0.07345807943594	1.28526542540785	1.10787316830863	3.42432052055345	1

Ratio observed/Expected

	LexoG0	NS-gDNA	NS-lexo	Cadoret	Karnani-BrIP	Karnani-Lexo	Karnani-Ori	Aladjem-K562	Aladjem-MCF7	Bubble	ORC	G4	CpG	ATtract	IR	Cruciform	Bends	Z-DNA
LexoG0	6.02631624415204	2.28219661881691	0.95313257473571	4.04427963946254	0.24010109588347	1.09360199566963	0.93888453048516	5.43105878224117	4.87443394536168	1.32026601563717	3.94759854505223	2.38074627668422	6.11904593473386	0.07241009697690	0.25019732987715	0.17070593069989	0.84975763427481	1.88895624254448
NS-gDNA	2.80587968725432	7.14562186946178	5.54391754890232	3.44176057901158	0.25212941182282	0.50121012605122	0.43457393684008	2.77766626998596	2.79500541939463	1.16565403466937	2.36152778375966	2.20354270568627	4.11877100212123	0.25144665572678	0.23799768258221	0.16421970442686	0.57069942626740	1.15008110666833
NS-lexo	1.04003463119592	5.77045052723548	13.1348643239624	0.28957635899994	0.33422869055322	0.13274115341227	0.27713636550194	0.04941948660234	0.02597335992392	0.61720469771517	0.17366387281872	1.70678383853811	0.02437316986730	0.26151037692089	0.22146897146803	0.14453240711088	0.26319372615680	0.47749390227514
Cadoret	3.90696826708957	2.7346422221758	0.29133355087229	3366.832062486	70.7445325994214	175.39729235707	137.208068373017	4.94837222670292	3.85340481915294	1.41035463758888	10.1361614626297	3.81345844604449	15.4850385083317	0.44326241134751	0.36285947739148	0.23521812891422	0.73861253241592	1.77267048862554
Karnani-BrIP	0.23817663272412	0.25586979193521	0.33624059561326	70.7411151095382	1056.89695936192	374.52368184575	782.152769978126	0.44377412899210	0.14125614050972	0.84348747018881	0.30860614586573	0.38811636193628	0.10137628963554	0.67361963190184	0.67032942183476	0.49296296601706	0.87274446139714	0.71317636135742
Karnani-Lexo	0.95437270059556	0.44508491916385	0.13354618628217	175.396688277989	385.344537522961	3226.6030053971	3152.03810017688	1.06405465521466	0.62058511383396	0.82907716740381	1.17715995097148	0.88136865357947	1.89276554844553	0.53125	0.86798884680959	0.56279119410626	0.99874513713306	1.1199014032516
Karnani-Ori	0.84562704951605	0.44104749271332	0.27882140532899	137.209709872451	804.763372470573	3152.08666580555	6572.56346247974	0.50477569078975	0.42609400480188	0.77477912771568	1.21985395056279	0.57501796819531	1.2057548377748	0.5	0.97331294439395	0.68836317658166	1.43403292725117	1.10678911663358
Aladjem-K562	6.97441510601577	3.17937079142902	0.05321330900830	5.43817884220217	0.44127405323761	1.13939824605029	0.50190170092940	64.6891574443458	17.4775517380961	1.68762583622449	3.728615240032	4.34647377305672	4.54277728542201	0.03341087517319	0.21373851272819	0.14401261689077	0.31656159629455	1.15436640929235
Aladjem-MCF7	6.77601657332457	3.25228262793334	0.02634399073558	4.63930735062958	0.14006058612785	0.66658046060209	0.42246219506587	16.7724705866232	33.4682922596019	1.72827593100206	3.6556028629845	4.37856460026424	5.28445439286146	0.04396605166864	0.18336358008636	0.12371494259922	0.42658084463233	1.34286621429908
Bubble	0.74865314739211	0.63715432134260	0.50398029086072	1.40527925172528	0.76912850463469	0.81030513821603	0.76613337951362	0.98100828506655	0.83415775159193	1.86957982628631	1.49062074703075	0.52172897965397	1.37505765512175	0.86718204974284	0.65525473204059	0.51010829784057	0.84479789918945	0.72280175638428
ORC	4.01697634110667	2.12704057251817	0.15540017132979	9.49117858832596	0.30824844884078	1.17574278636745	1.21836661510011	3.54753445995812	3.36462951916526	1.58049419624357	123.875502014354	3.0962386623251	19.0051675161437	0.26623256372178	0.44267948388047	0.31925765428660	0.49975912840426	2.27697164241238
G4	5.62576850038043	3.56543546276712	1.70790642301694	8.10978137469864	0.43156670457128	1.04032252972721	0.73017104862742	5.84012510863534	6.49951647063754	1.40940383510052	6.52356000522674	139.264944063553	12.0425813378176	0	0.10484002311298	0.06924735412402	0.32146863429134	0.47151766483955
CpG	6.81923520239088	3.91900068181887	0.02069551449137	15.4454216941315	0.10112181418491	1.88792961244735	1.20265565497059	4.22707342987737	4.87920160499316	1.7193778190315	19.0659570878422	5.23257431684169	65.1313509154102	0.02531094329949	0.18369452753786	0.13376174849999	0.19387308833505	4.58873088673063
ATtract	0.05818783872170	0.21498718411505	0.16231194903889	0.35276530364669	0.65138110376910	0.58330052892711	0.55534023539860	0.02657901536617	0.03778309421835	0.69774955426688	0.26592628057367	0	0.02385085593300	22.3605532120586	6.8933700004948	7.67886633459282	3.36022449977447	0.52370819946339
IR	0.26152273272136	0.26225793317579	0.24555633233431	0.40283429633343	0.73189738019118	1.07068347193221	1.30972863427529	0.21538892843370	0.18465128620986	0.91230555220637	0.48452197090887	0.10304493655909	0.18941638948130	10.050726052483	348.069691378975	268.906575008954	13.291376033013	5.46202710967023
Cruciform	0.17975967215960	0.18731015000063	0.16866652637619	0.42463491348393	0.67962098933971	1.0344727017809	1.58158556629437	0.14474804074807	0.12398815389928	0.89757602161535	0.39063056368578	0.06957507815040	0.13869242710952	13.9352164192384	298.287353739403	351.750159799511	33.3866525037545	6.11224516919308
Bends	0.93807882849339	0.61636303173121	0.27513951864433	0.73496890318325	0.86848110375648	0.99382167766082	1.42694166741651	0.32230950211156	0.43048580128054	0.93243479261253	0.50834024858820	0.30407515870155	0.20236575808448	2.08076769449002	8.53948376917059	10.3950381858969	103.757668763878	2.38754041847096
Z-DNA	3.37857555815478	2.01182110910681	0.50839427882638	2.6836285673621	0.84353351401521	1.46789060890814	1.56401606071246	1.35845096182639	1.57658814799349	1.20502346828338	3.07321703202777	0.47564307408374	6.98058677764614	0.65092324170173	5.52595367251958	5.05535774566165	3.62899746373739	361.490968725457

OverlapAnalyses - LexoG0

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
LexoG0	LexoG0	32665.2289764954	196851	0.16593885210893	1	0	6.02631624415203	1	1	1
LexoG0	NS-gDNA	27267.5892545402	62230	0.13851892677477	0.31612742632752	0	2.28219661881691	1.21439499561993	0.82345530375766	0.38390356451035
LexoG0	NS-lexo	13128.2891086475	12513	0.06669150326209	0.0635658442172	0.99999998889259	0.95313257473571	2.94550433182206	0.33950043433866	0.18723346949768
LexoG0	Cadoret	53.1615069100934	215	0.00027005962332	0.00109219663603	0	4.04427963946935	698.053191489362	0.00143255558772	0.76241134751773
LexoG0	Karnani-BrIP	162.431578483439	39	0.00082514987723	0.00019811938979	1	0.24010109588374	241.534969325153	0.00414018724822	0.04785276073619
LexoG0	Karnani-Lexo	57.6077953858564	63	0.00029264669920	0.00032003901428	0.21612130805891	1.09360199566789	615.159375	0.00162559499316	0.196875
LexoG0	Karnani-Ori	27.6924362430985	26	0.00014067714282	0.00013207959319	0.57778502247641	0.93888453048184	1312.34	0.00076199765304	0.17333333333333
LexoG0	Aladjem-K562	6764.97925600409	36741	0.03436598877325	0.18664370513738	0	5.43105878224112	3.12605802671071	0.31989169473358	0.58345905257975
LexoG0	Aladjem-MCF7	10784.0214041714	52566	0.05478266000259	0.26703445753387	0	4.87443394536169	2.0898243006529	0.47850912619189	0.55805509846594
LexoG0	Bubble	63231.1966007177	83482	0.32121348939409	0.42408725381125	0	1.32026601563717	1.59685740707692	0.62622999121162	0.67720687249541
LexoG0	ORC	1936.62043208056	7645	0.00983800149392	0.03883648038364	0	3.94759854505241	14.4711460707197	0.06910302716267	0.5620083805043
LexoG0	G4	30475.3180591127	72554	0.15481413891274	0.36857318479459	0	2.38074627668422	0.54845827865495	1	0.36857318479459
LexoG0	CpG	3900.11780505419	23865	0.01981253742706	0.12123382659981	0	6.11904593473386	6.90462995440196	0.14483035392251	0.83707471062785
LexoG0	ATtract	196851	14254	1	0.07241009697690	1	0.07241009697690	0.07718415139230	1	0.07241009697690
LexoG0	IR	12378.2296218771	3097	0.062881212805	0.01573271154324	1	0.25019732987715	1.35045895476311	0.74048899929388	0.02124638118628
LexoG0	Cruciform	14311.1606608134	2443	0.07270047223947	0.01241040177596	1	0.17070593069990	1.16406869064374	0.85905583410803	0.01444656014570
LexoG0	Bends	5500.3918899638	4674	0.02794190473995	0.02374384686895	1	0.84975763427481	3.59801502440094	0.27793102397244	0.08543071777156
LexoG0	Z-DNA	15816.6713061363	29877	0.08034844276196	0.15177469253394	0	1.88895624254446	1.05072378674979	0.95172490868728	0.15947327967205

OverlapAnalyses - NS-gDNA

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
NS-gDNA	LexoG0	27180.4241452094	76265	0.16767896053751	0.47048698935212	0	2.80587968725432	0.82345530375766	1	0.47048698935212
NS-gDNA	NS-gDNA	22684.9395281827	162098	0.13994583232478	1	0	7.14562186946175	1	1	1
NS-gDNA	NS-lexo	10900.9557712427	60434	0.06724916884380	0.37282384730224	0	5.54391754890228	2.42549116428005	0.41228762847166	0.9042809474645
NS-gDNA	Cadoret	44.1634438278405	152	0.00027244903594	0.00093770435168	0	3.44176057901037	574.815602836879	0.00173968833668	0.53900709219858
NS-gDNA	Karnani-BrIP	134.85138347911	34	0.00083191269157	0.00020974965761	1	0.25212941182221	198.893251533742	0.00502782267517	0.04171779141104
NS-gDNA	Karnani-Lexo	47.8841083861532	24	0.00029540221585	0.00014805858184	0.99989486582404	0.50121012604967	506.55625	0.00197411442460	0.075
NS-gDNA	Karnani-Ori	23.0110440417794	10	0.00014195760615	0.00006169107576	0.99803645262654	0.43457393683849	1080.65333333333	0.00092536613653	0.06666666666667
NS-gDNA	Aladjem-K562	5670.58763329389	15751	0.03498246513401	0.09716961344371	0	2.77766626998597	2.57416906194915	0.38847487322483	0.25013101268838
NS-gDNA	Aladjem-MCF7	9027.88947202306	25233	0.05569402134525	0.15566509148786	0	2.79500541939461	1.72087690429428	0.5810990882059	0.26788046074632
NS-gDNA	Bubble	52132.1062619035	60768	0.32160857173995	0.37488432923293	0	1.16565403466937	1.3149407012022	0.76049056743451	0.49295066275127
NS-gDNA	ORC	1615.05616246777	3814	0.00996345520899	0.02352897629828	0	2.36152778375973	11.916341983386	0.08391837036854	0.28037932808939
NS-gDNA	G4	25685.9101727156	56600	0.15845914306601	0.34917148885242	0	2.20354270568626	0.45163087844822	1	0.34917148885242
NS-gDNA	CpG	3254.61162883204	13405	0.02007804925928	0.08269688706831	0	4.11877100212125	5.68565415643634	0.17588125701736	0.47018589968432
NS-gDNA	ATtract	162098	40759	1	0.25144665572678	1	0.25144665572678	0.06355769883003	1	0.25144665572678
NS-gDNA	IR	10432.8746946614	2483	0.06436152632766	0.01531789411343	1	0.23799768258221	1.11204258880672	0.89924613505410	0.01703415062497
NS-gDNA	Cruciform	12063.1078159218	1981	0.07441860982814	0.01222100210983	1	0.16421970442685	0.95855853724882	1	0.01222100210983
NS-gDNA	Bends	4617.14149115921	2635	0.02848364255672	0.01625559846512	1	0.57069942626740	2.96280455484272	0.33751804463966	0.04816216117417
NS-gDNA	Z-DNA	13332.9727017437	15334	0.0822525429169	0.09459709558415	0	1.15008110666833	0.86522407498345	1	0.09459709558415

OverlapAnalyses - NSlexo

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
NS-lexo	LexoG0	12972.6449440302	13492	0.19411119007691	0.20188235998264	0.00000020965063	1.04003463119592	0.33950043433866	1	0.20188235998264
NS-lexo	NS-gDNA	10806.2619557496	62357	0.16169535029776	0.93305501937723	0	5.77045052723549	0.41228762847166	1	0.93305501937723
NS-lexo	NS-lexo	5088.0616922763	66831	0.07613325690587	1	0	13.1348643239625	1	1	1
NS-lexo	Cadoret	20.7199234797608	6	0.00031003461686	0.00008977869551	0.9998486229073	0.28957635899866	236.989361702128	0.00421959868923	0.02127659574468
NS-lexo	Karnani-BrIP	62.8312308114726	21	0.00094015098998	0.00031422543430	0.99999999914462	0.33422869055376	82.001226993865	0.01219493947419	0.025766871165644
NS-lexo	Karnani-Lexo	22.6003761674272	3	0.00033817204841	0.00004488934775	0.99999966385382	0.13274115341158	208.846875	0.00478819709416	0.009375
NS-lexo	Karnani-Ori	10.8249958267836	3	0.00016197566738	0.00004488934775	0.99439542983815	0.27713636550115	445.54	0.00224446738788	0.02
NS-lexo	Aladjem-K562	2913.83035114607	144	0.04359998131325	0.00215468869237	1	0.04941948660235	1.06129805783615	0.94224237255166	0.00228676692445
NS-lexo	Aladjem-MCF7	4581.61748609277	119	0.06855527354211	0.00178061079439	1	0.02597335992391	0.70949625776315	1	0.00178061079439
NS-lexo	Bubble	22473.9054180062	13871	0.33627965192809	0.20755338091604	1	0.61720469771517	0.54213378327952	1	0.20755338091604
NS-lexo	ORC	788.880253424829	137	0.01180410667841	0.00204994688093	1	0.17366387281875	4.91296037638756	0.20354326584968	0.01007130779975
NS-lexo	G4	13896.3115682622	23718	0.20793212084604	0.35489518337298	0	1.70678383853811	0.18620182381999	1	0.35489518337298
NS-lexo	CpG	1600.12014080708	39	0.02394278315163	0.00058356152085	1	0.02437316986731	2.3441248684672	0.42659843485807	0.00136794107330
NS-lexo	ATtract	66831	17477	1	0.26151037692089	1	0.26151037692089	0.02620405292174	1	0.26151037692089
NS-lexo	IR	5644.13150841947	1250	0.08445379402402	0.01870389489907	1	0.22146897146803	0.45848140169861	1	0.01870389489907
NS-lexo	Cruciform	6531.40716929817	944	0.09773020259008	0.01412518142778	1	0.14453240711089	0.39520182607358	1	0.01412518142778
NS-lexo	Bends	2405.07252677776	633	0.03598737901240	0.00947165237689	1	0.26319372615680	1.22152766354115	0.81864703505858	0.01156988539781
NS-lexo	Z-DNA	7223.12888932459	3449	0.10808051487071	0.05160778680552	1	0.47749390227515	0.35672118197151	1	0.05160778680552

OverlapAnalyses - Cadoret

p-values here should not be used or considered. The p-value is calculated here assuming genome-wide - not 1% ENCODE.

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
Cadoret	LexoG0	52.2143990055918	204	0.18515744328224	0.72340425531914	0	3.90696826708956	0.00143255558772	1	0.72340425531914
Cadoret	NS-gDNA	43.5157473379895	119	0.15431116077301	0.42198581560283	0	2.7346422221758	0.00173968833668	1	0.42198581560283
Cadoret	NS-lexo	20.5949502967825	6	0.07303173864107	0.02127659574468	0.99988972470178	0.29133355087229	0.00421959868923	1	0.02127659574468
Cadoret	Cadoret	0.08375826140614	282	0.00029701511136	1	0	3366.83206248264	1	1	1
Cadoret	Karnani-BrlP	0.25443662341978	18	0.00090225752985	0.06382978723404	0	70.7445325993923	0.34601226993865	1	0.06382978723404
Cadoret	Karnani-Lexo	0.09122147659747	16	0.00032348041346	0.05673758865248	0	175.397292356961	0.88125	1	0.05673758865248
Cadoret	Karnani-Ori	0.04372920682537	6	0.00015506810931	0.02127659574468	0.00000000000005	137.208068373163	1.88	0.53191489361702	0.04
Cadoret	Aladjem-K562	11.5189394387938	57	0.04084730297444	0.20212765957446	0	4.94837222670291	0.00447825189372	1	0.20212765957446
Cadoret	Aladjem-MCF7	18.1657529601023	70	0.06441756368830	0.24822695035461	0	3.85340481915294	0.00299378947927	1	0.24822695035461
Cadoret	Bubble	92.8844394938533	131	0.32937744501366	0.46453900709219	0.00000092779136	1.41035463758888	0.00228758700131	1	0.46453900709219
Cadoret	ORC	3.15701364051645	32	0.01119508383161	0.11347517730496	0	10.13616146263	0.02073072116444	1	0.11347517730496
Cadoret	G4	54.2814358485302	207	0.19248736116500	0.73404255319148	0	3.8134584460445	0.00078569697172	1	0.73404255319148
Cadoret	CpG	6.39326792417929	99	0.0226711628517	0.35106382978723	0	15.485038508332	0.00989126622237	1	0.35106382978723
Cadoret	ATtract	282	125	1	0.44326241134751	1	0.44326241134751	0.00011057058736	1	0.44326241134751
Cadoret	IR	22.0471022488104	8	0.07818121364826	0.02836879432624	0.99961941109732	0.36285947739148	0.00193460752164	1	0.02836879432624
Cadoret	Cruciform	25.5082379393813	6	0.09045474446589	0.02127659574468	0.99999802032802	0.23521812891422	0.00166759310728	1	0.02127659574468
Cadoret	Bends	9.4772288483973	7	0.03360719449786	0.02482269503546	0.73338816051831	0.73861253241592	0.00515435652793	1	0.02482269503546
Cadoret	Z-DNA	28.2060317023543	50	0.10002138901544	0.17730496453900	0.00002506008660	1.77267048862555	0.00150522023186	1	0.17730496453900

OverlapAnalyses - Karnani

p-values here should not be used or considered. The p-value is calculated here assuming genome-wide - not 1% ENCODE.

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
Karnani-BrIP	LexoG0	159.545458197888	38	0.19576129840231	0.04662576687116	1	0.23817663272412	0.00414018724822	1	0.04662576687116
Karnani-BrIP	NS-gDNA	132.880086167456	34	0.16304305051221	0.04171779141104	1	0.25586979193521	0.00502782267517	1	0.04171779141104
Karnani-BrIP	NS-lexo	62.4552783749927	21	0.07663224340489	0.02576687116564	0.99999999964079	0.33624059561326	0.01219493947419	1	0.02576687116564
Karnani-BrIP	Cadoret	0.25444891520499	18	0.00031220725792	0.02208588957055	0	70.7411151094852	2.89007092198582	0.34601226993865	0.06382978723404
Karnani-BrIP	Karnani-BrIP	0.77112531432778	815	0.00946166029850	1	0	1056.89695936187	1	1	1
Karnani-BrIP	Karnani-Lexo	0.27768604507884	104	0.00034071907371	0.12760736196319	0	374.523681845342	2.546875	0.39263803680981	0.325
Karnani-BrIP	Karnani-Ori	0.13296635132131	104	0.00016314889732	0.12760736196319	0	782.152769979273	5.43333333333333	0.18404907975460	0.69333333333333
Karnani-BrIP	Aladjem-K562	36.0543775644132	16	0.04423850007903	0.01963190184049	0.99988579465140	0.44377412899210	0.01294246557939	1	0.01963190184049
Karnani-BrIP	Aladjem-MCF7	56.6347060816736	8	0.06949043691002	0.00981595092024	1	0.14125614050972	0.00865226392059	1	0.00981595092024
Karnani-BrIP	Bubble	273.862989272727	231	0.33602820769659	0.28343558282208	0.99926844119917	0.84348747018881	0.00661128867401	1	0.28343558282208
Karnani-BrIP	ORC	9.72112850048434	3	0.01192776503126	0.00368098159509	0.98770088641200	0.30860614586573	0.05991325442917	1	0.00368098159509
Karnani-BrIP	G4	172.628640714198	67	0.21181428308490	0.08220858895705	1	0.38811636193628	0.00227071997147	1	0.08220858895705
Karnani-BrIP	CpG	19.7284789884283	2	0.02420672268518	0.00245398773006	0.99999952194310	0.10137628963556	0.02858646089091	1	0.00245398773006
Karnani-BrIP	ATtract	815	549	1	0.67361963190184	1	0.67361963190184	0.00031955683935	1	0.67361963190184
Karnani-BrIP	IR	70.1147800902968	47	0.08603040501876	0.05766871165644	0.99849810247631	0.67032942183476	0.00559115294375	1	0.05766871165644
Karnani-BrIP	Cruciform	81.1419979946637	40	0.09956073373578	0.04907975460122	0.99999989942580	0.49296296601706	0.00481946234905	1	0.04907975460122
Karnani-BrIP	Bends	29.7910799208943	26	0.03655347229557	0.03190184049079	0.72446159412399	0.87274446139713	0.01489645592294	1	0.03190184049079
Karnani-BrIP	Z-DNA	89.7393736917831	64	0.11010966097151	0.07852760736196	0.99836792980757	0.71317636135742	0.00435019322330	1	0.07852760736196
Karnani-Lexo	LexoG0	56.5816687404218	54	0.17681771481381	0.16875	0.61429414795394	0.95437270059556	0.00162559499316	1	0.16875
Karnani-Lexo	NS-gDNA	47.1820075131979	21	0.14744377347874	0.065625	0.99999551184755	0.44508491916385	0.00197411442460	1	0.065625
Karnani-Lexo	NS-lexo	22.4641383143753	3	0.07020043223242	0.009375	0.99999979725865	0.13354618628217	0.00478819709416	1	0.009375
Karnani-Lexo	Cadoret	0.09122179077081	16	0.00028506809615	0.05	0	175.396688278109	1.13475177304965	0.88125	0.05673758865248
Karnani-Lexo	Karnani-BrIP	0.27767358709108	107	0.00086772995966	0.334375	0	385.344537523121	0.39263803680981	1	0.334375
Karnani-Lexo	Karnani-Lexo	0.09917551042517	320	0.00030992347007	1	0	3226.60300540044	1	1	1
Karnani-Lexo	Karnani-Ori	0.04758825725857	150	0.00014871330393	0.46875	0	3152.0381001757	2.13333333333333	0.46875	1
Karnani-Lexo	Aladjem-K562	12.2174175323516	13	0.03817942978859	0.040625	0.33994607153123	1.06405465521467	0.00508170427657	1	0.040625
Karnani-Lexo	Aladjem-MCF7	19.3365901509693	12	0.06042684422177	0.0375	0.95268803219380	0.62058511383396	0.00339720791974	1	0.0375
Karnani-Lexo	Bubble	103.729789434803	86	0.32415559198375	0.26875	0.98139220062551	0.82907716740381	0.00259584340574	1	0.26875
Karnani-Lexo	ORC	3.39800890838923	4	0.01061877783871	0.0125	0.25484792734925	1.17715995097144	0.02352422259795	1	0.0125
Karnani-Lexo	G4	56.7299503980961	50	0.17728109499405	0.15625	0.81844106122754	0.88136865357946	0.00089157103174	1	0.15625
Karnani-Lexo	CpG	6.86825687982156	13	0.02146330274944	0.040625	0.01022223586066	1.8927655484455	0.01122413188355	1	0.040625
Karnani-Lexo	ATtract	320	170	1	0.53125	1	0.53125	0.00012547017005	1	0.53125
Karnani-Lexo	IR	23.0417707249493	20	0.07200553351546	0.0625	0.70115218149244	0.86798884680960	0.00219529931534	1	0.0625
Karnani-Lexo	Cruciform	26.6528690517636	15	0.08329021578676	0.046875	0.99184630002608	0.56279119410626	0.00189230423521	1	0.046875
Karnani-Lexo	Bends	10.0125643952624	10	0.03128926373519	0.03125	0.41854683961788	0.99874513713306	0.00584891520900	1	0.03125
Karnani-Lexo	Z-DNA	29.4668797665451	33	0.09208399927045	0.103125	0.21467112962695	1.1199014032516	0.00170805132694	1	0.103125
Karnani-Ori	LexoG0	27.1987515219181	23	0.18132501014612	0.15333333333333	0.78085457782465	0.84562704951605	0.00076199765304	1	0.15333333333333
Karnani-Ori	NS-gDNA	22.6732952011135	10	0.15115530134075	0.06666666666666	0.99873415897237	0.44104749271331	0.00092536613653	1	0.06666666666667
Karnani-Ori	NS-lexo	10.7595756375308	3	0.07173050425020	0.02	0.99527434127228	0.27882140532899	0.00224446738788	1	0.02
Karnani-Ori	Cadoret	0.04372868367386	6	0.00029152455782	0.04	0.00000000000005	137.209709872576	0.53191489361702	1	0.04
Karnani-Ori	Karnani-BrIP	0.13295833739487	107	0.00088638891596	0.71333333333333	0	804.7633724707	0.18404907975460	1	0.71333333333333
Karnani-Ori	Karnani-Lexo	0.0475875240447	150	0.00031725016029	1	0	3152.08666580555	0.46875	1	1
Karnani-Ori	Karnani-Ori	0.02282214555343	150	0.00015214763702	1	0	6572.56346248343	1	1	1
Karnani-Ori	Aladjem-K562	5.94323390515559	3	0.03962155936770	0.02	0.84907816548504	0.50477569078975	0.00238204887964	1	0.02
Karnani-Ori	Aladjem-MCF7	9.38759981347283	4	0.06258399875648	0.02666666666666	0.96132914007649	0.42609400480187	0.00159244121237	1	0.02666666666667
Karnani-Ori	Bubble	49.0462360699328	38	0.32697490713288	0.25333333333333	0.96881237242629	0.77477912771568	0.00121680159644	1	0.25333333333333
Karnani-Ori	ORC	1.63954053604306	2	0.01093027024028	0.01333333333333	0.22638718712149	1.21985395056282	0.01102697934279	1	0.01333333333333
Karnani-Ori	G4	27.8252174453187	16	0.18550144963545	0.10666666666666	0.99395208621649	0.57501796819531	0.00041792392112	1	0.10666666666667
Karnani-Ori	CpG	3.31742396935516	4	0.02211615979570	0.02666666666667	0.23917837857349	1.20575483777478	0.00526131182041	1	0.02666666666667
Karnani-Ori	ATtract	150	75	1	0.5	1	0.5	0.00005881414221	1	0.5
Karnani-Ori	IR	11.3016066038752	11	0.07534404402583	0.07333333333333	0.45806231699949	0.97331294439395	0.00102904655406	1	0.07333333333333
Karnani-Ori	Cruciform	13.0744936774408	9	0.08716329118293	0.06	0.85090514499781	0.68836317658166	0.00088701761025	1	0.06
Karnani-Ori	Bends	4.88133840372691	7	0.03254225602484	0.04666666666666	0.11777268054219	1.43403292725116	0.00274167900422	1	0.04666666666667

OverlapAnalyses - Aladjem

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
Aladjem-K562	LexoG0	6682.42415909547	46606	0.106119073209818	0.740118467230948	0	6.97441510601574	0.31989169473358	1	0.740118467230948
Aladjem-K562	NS-gDNA	5619.350862807	17866	0.089237122847136	0.283717901891347	0	3.17937079142901	0.38847487322483	1	0.283717901891347
Aladjem-K562	NS-lexo	2912.8051400783	155	0.046256294803613	0.002461450508964	1	0.05321330900831	0.94224237255166	1	0.002461450508964
Aladjem-K562	Cadoret	11.5847606024175	63	0.000183969773420	0.00100046052945	0	5.43817884219838	223.301418439716	0.004478251893725	0.223404255319149
Aladjem-K562	Karnani-BrIP	36.258646713148	16	0.000575799125203	0.000254085213825	0.999868183098518	0.44127405323702	77.2650306748466	0.012942465579394	0.019631901840491
Aladjem-K562	Karnani-Lexo	12.2871875997001	14	0.000195124543039	0.0002223245621	0.25454340042855	1.13939824605117	196.784375	0.005081704276572	0.04375
Aladjem-K562	Karnani-Ori	5.97726605516514	3	0.000094920932734	0.000047640977593	0.846768646265504	0.50190170093024	419.806666666667	0.002382048879643	0.02
Aladjem-K562	Aladjem-K562	973.439792505815	62971	0.015458541114256	1	0	64.6891574443458	1	1	1
Aladjem-K562	Aladjem-MCF7	1666.13725059251	29120	0.026458802474036	0.46243508916803	0	17.4775517380962	0.668517437231270	1	0.46243508916803
Aladjem-K562	Bubble	17701.1985469664	29873	0.281100801114266	0.474392974543838	0	1.68762583622449	0.510821422197706	1	0.474392974543838
Aladjem-K562	ORC	360.455534690227	1344	0.005724151350467	0.021343157961600	0	3.72861524003238	4.62919944129971	0.216020072731892	0.098801734911417
Aladjem-K562	G4	2994.15127744983	13014	0.047548097972874	0.206666560797828	0	4.34647377305673	0.17544724824959	1	0.206666560797828
Aladjem-K562	CpG	705.295417030837	3204	0.011200321053038	0.050880564069175	0	4.54277728542211	2.20873377762189	0.452748090390815	0.112381620484041
Aladjem-K562	ATtract	20801.6101463178	695	0.330336347625375	0.011036826475675	1	0.03341087517319	0.024690568995458	1	0.011036826475675
Aladjem-K562	IR	1216.43964244599	260	0.019317457916275	0.004128884724715	1	0.21373851272817	0.432000603707312	1	0.004128884724715
Aladjem-K562	Cruciform	1395.7110449043	201	0.022164346205464	0.003191945498722	1	0.14401261689075	0.372375906236325	1	0.003191945498722
Aladjem-K562	Bends	726.556861894228	230	0.011537959725814	0.003652474948780	1	0.31656159629455	1.1509751238325	0.868828508360993	0.004203907806474
Aladjem-K562	Z-DNA	1534.17492552097	1771	0.024363197750090	0.028124057105652	0.000000001028341	1.15436640929235	0.336117812840270	1	0.028124057105652
Aladjem-MCF7	LexoG0	10682.8251106954	72387	0.113411806472694	0.768480280269653	0	6.77601657332458	0.47850912619189	1	0.768480280269653
Aladjem-MCF7	NS-gDNA	8971.85249196557	29179	0.095247651063910	0.30977228090663	0	3.25228262793333	0.5810990882059	1	0.30977228090663
Aladjem-MCF7	NS-lexo	4593.07783754157	121	0.048761376267755	0.001284569244652	1	0.02634399073558	1.4094507040146	0.70949625776315	0.001810537026230
Aladjem-MCF7	Cadoret	18.3217005418614	85	0.000194508206825	0.000902383353688	0	4.63930735063551	334.024822695035	0.002993789479272	0.301418439716312
Aladjem-MCF7	Karnani-BrIP	57.1181388084246	8	0.000606381854755	0.000084930197994	1	0.14006058612715	115.576687116564	0.008652263920590	0.009815950920245
Aladjem-MCF7	Karnani-Lexo	19.502521853396	13	0.000207044130290	0.000138011571735	0.919179801473310	0.66658046060521	294.359375	0.00339720791974	0.040625
Aladjem-MCF7	Karnani-Ori	9.468302836886	4	0.000100518104325	0.000042465098997	0.958931796676867	0.42246219506383	627.966666666667	0.001592441212375	0.026666666666667
Aladjem-MCF7	Aladjem-K562	1670.89277964523	28025	0.017738656825153	0.297521099846064	0	16.7724705866228	1.49584729478649	0.668517437231270	0.445046132346636
Aladjem-MCF7	Aladjem-MCF7	2814.45492555645	94195	0.029879026758920	1	0	33.4682922596021	1	1	1
Aladjem-MCF7	Bubble	26966.1800896447	46605	0.286280376767815	0.494771484686024	0	1.72827593100206	0.764110842513425	1	0.494771484686024
Aladjem-MCF7	ORC	586.223416580415	2143	0.006223508854827	0.022750676787515	0	3.65560286298464	6.9245754612953	0.144413185413236	0.157538778210689
Aladjem-MCF7	G4	5692.04802836437	24923	0.060428345754704	0.264589415574075	0	4.37856460026422	0.26244229167189	1	0.264589415574075
Aladjem-MCF7	CpG	1153.38302630322	6095	0.012244631098288	0.064706194596310	0	5.28445439286157	3.30392844615924	0.302669993099422	0.213784636969484
Aladjem-MCF7	ATtract	39735.2032692531	1747	0.421839835121324	0.018546631986830	1	0.04396605166864	0.036933320838596	1	0.018546631986836
Aladjem-MCF7	IR	2312.34577662759	424	0.024548498079805	0.004501300493657	1	0.18336358008635	0.646206934401712	1	0.004501300493657
Aladjem-MCF7	Cruciform	2659.33922845387	329	0.028232275900567	0.003492754392484	1	0.12371494259921	0.557017491987274	1	0.003492754392484
Aladjem-MCF7	Bends	1272.91229044277	543	0.013513586606962	0.005764637188810	1	0.42658084463236	1.72168302535139	0.580827007802962	0.009924877995284
Aladjem-MCF7	Z-DNA	2928.06532633803	3932	0.031085145987985	0.041743192313817	0	1.3428662142991	0.502780921066678	1	0.041743192313817

OverlapAnalyses - Bubble-seq

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
Bubble	LexoG0	62804.785051379	47019	0.50947308476547	0.38141862842124	1	0.74865314739211	0.62622999121162	1	0.38141862842124
Bubble	NS-gDNA	51946.5989499316	33098	0.42139136354731	0.26849132826062	1	0.63715432134260	0.76049056743451	1	0.26849132826062
Bubble	NS-lexo	22590.1691126769	11385	0.18325169226825	0.09235524117007	1	0.50398029086073	1.84456315183074	0.54213378327952	0.17035507474076
Bubble	Cadoret	93.9315085153806	132	0.00076197339678	0.00107078540487	0.00008390237959	1.40527925172612	437.141843971631	0.00228758700131	0.46808510638297
Bubble	Karnani-BrIP	276.936817081224	213	0.00224651440759	0.00172785826695	0.99996363498984	0.76912850463479	151.256441717791	0.00661128867401	0.26134969325153
Bubble	Karnani-Lexo	104.898754791455	85	0.00085093981530	0.00068952090465	0.97389467125740	0.81030513821622	385.23125	0.00259584340574	0.265625
Bubble	Karnani-Ori	49.5997185556814	38	0.00040235344481	0.00030825640443	0.94702961832874	0.76613337951384	821.826666666667	0.00121680159644	0.25333333333333
Bubble	Aladjem-K562	17799.0341833	17461	0.14438595472930	0.14164381783668	0.99694137198677	0.98100828506654	1.95763129059408	0.51082142219770	0.27728636991631
Bubble	Aladjem-MCF7	27038.0512042924	22554	0.21933295913406	0.18295828804127	1	0.83415775159193	1.30871065343171	0.76411084251342	0.23943946069324
Bubble	Bubble	65936.7405802985	123274	0.53487954134934	1	0	1.86957982628631	1	1	1
Bubble	ORC	4140.55688698441	6172	0.03358824153499	0.05006732968833	0	1.49062074703075	9.06226567668897	0.11034768077615	0.45372344335808
Bubble	G4	96421.7092816364	50306	0.78217393190483	0.40808280740464	1	0.52172897965397	0.34346102302203	1	0.40808280740464
Bubble	CpG	8579.27662600849	11797	0.06959518329906	0.09569738955497	0	1.37505765512174	4.32388635566468	0.23127342343073	0.41378463696948
Bubble	ATtract	123274	106901	1	0.86718204974284	1	0.86718204974284	0.04833503044808	1	0.86718204974284
Bubble	IR	39160.3429098322	25660	0.31766911846644	0.20815419309830	1	0.65525473204059	0.84569789937296	1	0.20815419309830
Bubble	Cruciform	45400.1632556035	23159	0.36828660752148	0.18786605448026	1	0.51010829784057	0.72897472591155	1	0.18786605448026
Bubble	Bends	15229.6780239916	12866	0.12354331021944	0.10436912893229	1	0.84479789918946	2.2531849171099	0.44381621428687	0.23516294712215
Bubble	Z-DNA	50273.8125344016	36338	0.40782170234114	0.29477424274380	1	0.72280175638428	0.65799474774217	1	0.29477424274380

OverlapAnalyses - ORC

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
ORC	LexoG0	1904.41748977103	7650	0.13999981546504	0.56237594648239	0	4.01697634110667	0.069103027162676	1	0.562375946482394
ORC	NS-gDNA	1593.2935383493	3389	0.11712809956254	0.24913621995148	0	2.12704057251817	0.083918370368542	1	0.24913621995148
ORC	NS-lexo	785.069919524683	122	0.05771299856830	0.00896860986547	1	0.15540017132979	0.20354326584968	1	0.00896860986547
ORC	Cadoret	3.16082978744761	30	0.00023236269848	0.00220539586855	0	9.49117858833682	48.2375886524823	0.02073072116444	0.10638297872340
ORC	Karnani-BrIP	9.73240907224702	3	0.00071546049196	0.00022053958685	0.98745992350913	0.30824844884036	16.6907975460123	0.05991325442917	0.00368098159509
ORC	Karnani-Lexo	3.40210464940111	4	0.00025009958460	0.00029405278247	0.25619561850016	1.1757427863673	42.509375	0.02352422259795	0.0125
ORC	Karnani-Ori	1.64154202455618	2	0.00012067499996	0.00014702639123	0.22742500980487	1.21836661509822	90.6866666666667	0.01102697934279	0.013333333333333
ORC	Aladjem-K562	358.84077078564	1273	0.02637953177869	0.09358229802249	0	3.54753445995814	0.21602007273189	1	0.09358229802249
ORC	Aladjem-MCF7	581.936284172455	1958	0.04277999589593	0.14393883702124	0	3.36462951916528	0.144413185413238	1	0.14393883702124
ORC	Bubble	4099.35070650618	6479	0.30135637039669	0.47629199441299	0	1.58049419624357	0.11034768077615	1	0.47629199441299
ORC	ORC	109.81186577491	13603	0.00807262116995	1	0	123.875502014356	1	1	1
ORC	G4	1496.0087077143	4632	0.10997638077735	0.34051312210541	0	3.09623866232509	0.037900127327488	1	0.34051312210541
ORC	CpG	219.361391919259	4169	0.01612595691533	0.306476512534	0	19.0051675161438	0.477130831287268	1	0.306476512534
ORC	ATtract	10528.3890175402	2803	0.77397552139529	0.20605748731897	1	0.26623256372178	0.00533365851019	1	0.20605748731897
ORC	IR	607.663128279579	269	0.04467125841943	0.01977504962140	1	0.44267948388047	0.093320801833075	1	0.01977504962140
ORC	Cruciform	701.627657136453	224	0.05157889121050	0.01646695581856	1	0.31925765428661	0.080440670348775	1	0.01646695581856
ORC	Bends	286.137844958637	143	0.02103490737033	0.01051238697346	1	0.49975912840425	0.248633729962890	1	0.01051238697346
ORC	Z-DNA	774.713205532548	1764	0.05695164342663	0.12967727707123	0	2.27697164241237	0.072608194376241	1	0.12967727707123

OverlapAnalyses - G4

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
G4	LexoG0	30937.9954024469	174050	0.08619818900315	0.48493105648381	0	5.62576850038042	1.82329274425835	0.54845827865495	0.884171276752468
G4	NS-gDNA	26159.497479058	93270	0.07288453174148	0.25986509415825	0	3.5654354627671	2.21419758417747	0.45163087844822	0.57539266369727
G4	NS-lexo	14276.5432996783	24383	0.03977672637316	0.06793492645932	0	1.70790642301694	5.37051667639269	0.18620182381999	0.36484565545929
G4	Cadoret	56.1050882848822	455	0.00015631772327	0.00126770256075	0	8.1097813747243	1272.75531914894	0.00078569697172	1.61347517730496
G4	Karnani-BrIP	178.419695459124	77	0.00049710572488	0.00021453427951	1	0.43156670457180	440.388957055215	0.00227071997147	0.094478527607362
G4	Karnani-Lexo	58.6356617845314	61	0.00016336830460	0.00016995572792	0.34725744767800	1.0403225297287	1121.615625	0.00089157103174	0.190625
G4	Karnani-Ori	28.7603843503231	21	0.00008013101733	0.00005850934895	0.91698818384550	0.73017104862731	2392.78	0.00041792392112	0.14
G4	Aladjem-K562	3077.16010628405	17971	0.00857345878374	0.05007007191077	0	5.84012510863518	5.6997189182322	0.17544724824959	0.285385336107097
G4	Aladjem-MCF7	5833.20315769293	37913	0.01625223424271	0.10563166414519	0	6.49951647063752	3.81036148415521	0.26244229167189	0.40249482456606
G4	Bubble	98550.1788350772	138897	0.27457651444505	0.3869891924874	0	1.40940383510052	2.91153852393854	0.34346102302203	1.12673394227493
G4	ORC	1544.40213501962	10075	0.00430295064045	0.02807055670252	0	6.52356000522625	26.3851356318459	0.03790012732748	0.740645445857531
G4	G4	2577.22431451379	358917	0.00718055794101	1	0	139.264944063556	1	1	1
G4	CpG	2942.9736869375	35441	0.00819959402017	0.09874427792498	0	12.042581337817	12.5891616976499	0.07943340660932	1.24310768151526
G4	ATtract	15534.7185571998	0	0.04328220328711	0	1	0	0.14072930320533	1	0
G4	IR	1049.21762447152	110	0.00292328762491	0.00030647754216	1	0.10484002311284	2.4622820136383	0.40612732191565	0.000754634139648
G4	Cruciform	1126.39682752648	78	0.00313832119271	0.00021732043898	1	0.06924735412411	2.12243799746904	0.47115628404338	0.000461249157333
G4	Bends	1975.30935296281	635	0.00550352686822	0.00176921126611	1	0.32146863429140	6.56023468772276	0.15243357099273	0.011606441117874
G4	Z-DNA	1177.05028122109	555	0.00327944979262	0.00154631850817	1	0.47151766483946	1.91577705660055	0.52198140517166	0.002962401520166

OverlapAnalyses - CpG

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
CpG	LexoG0	3840.45999627903	26189	0.13470571716166	0.91858996843212	0	6.8192352023909	0.14483035392251	1	0.91858996843212
CpG	NS-gDNA	3215.10533500397	12600	0.11277114468621	0.44195019291476	0	3.91900068181885	0.17588125701736	1	0.44195019291476
CpG	NS-lexo	1594.54842322259	33	0.05592944311548	0.00115748860049	1	0.02069551449137	0.42659843485807	1	0.00115748860049
CpG	Cadoret	6.40966636977646	99	0.00022482168957	0.00347246580147	0	15.4454216941486	101.099290780142	0.00989126622237	0.35106382978723
CpG	Karnani-BrIP	19.7781261750107	2	0.00069372592686	0.00007015082427	0.99999944635117	0.10112181418515	34.9815950920245	0.02858646089091	0.00245398773006
CpG	Karnani-Lexo	6.88584993543077	13	0.00024152402439	0.00045598035776	0.01125989151076	1.88792961245194	89.09375	0.01122413188355	0.040625
CpG	Karnani-Ori	3.32597280316565	4	0.00011665986682	0.00014030164854	0.24214888736337	1.20265565496892	190.066666666667	0.00526131182041	0.026666666666667
CpG	Aladjem-K562	703.086910909465	2972	0.02466106316764	0.10424412486846	0	4.22707342987743	0.45274809039081	1	0.10424412486846
CpG	Aladjem-MCF7	1146.49904899919	5594	0.04021392665728	0.19621185548930	0	4.87920160499316	0.30266999309942	1	0.19621185548930
CpG	Bubble	8505.40226710468	14624	0.29833048990195	0.51294282707821	0	1.7193778190315	0.23127342343073	1	0.51294282707821
CpG	ORC	219.658524390089	4188	0.00770461327218	0.14689582602595	0	19.0659570878414	2.09586120708667	0.47713083128726	0.30787326325075
CpG	G4	2854.61784115009	14937	0.10012689726938	0.52392143107681	0	5.23257431684168	0.07943340660932	1	0.52392143107681
CpG	CpG	437.730825467242	28510	0.01535358910793	1	0	65.1313509154122	1	1	1
CpG	ATtract	20070.3701157653	508	0.70397650353438	0.01781830936513	1	0.02531094329949	0.01117860796335	1	0.01781830936513
CpG	IR	1159.53372620806	213	0.04067112333244	0.00747106278498	1	0.18369452753785	0.19558744837616	1	0.00747106278498
CpG	Cruciform	1338.2002105035	179	0.04693792390401	0.00627849877236	1	0.13376174849998	0.16859248045604	1	0.00627849877236
CpG	Bends	557.065454145723	108	0.01953930039094	0.00378814451069	1	0.19387308833505	0.52110178940249	1	0.00378814451069
CpG	Z-DNA	1477.09686344872	6778	0.05180978125039	0.23774114345843	0	4.58873088673059	0.15217669790977	1	0.23774114345843

OverlapAnalyses - OtherNonB

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
ATtract	LexoG0	274713.760661426	15985	0.10771369458342	0.00626762708853	1	0.05818783872170	12.9560276554348	0.07718415139230	0.08120354989306
ATtract	NS-gDNA	232325.476542218	49947	0.09109349078096	0.01958393307421	1	0.21498718411505	15.7337351478735	0.06355769883003	0.30812841614332
ATtract	NS-lexo	127002.356401133	20614	0.04979689767207	0.00808263151724	1	0.16231194903889	38.1620355822897	0.02620405292174	0.30844967155960
ATtract	Cadoret	498.915279308164	176	0.00019562182793	0.00006900859353	1	0.35276530364845	9043.99645390071	0.00011057058736	0.62411347517730
ATtract	Karnani-BrIP	1587.39637059917	1034	0.00062240903926	0.00040542548699	1	0.65138110376913	3129.33374233129	0.00031955683935	1.26871165644172
ATtract	Karnani-Lexo	521.172165844461	304	0.00020434862586	0.00011919666155	1	0.58330052892871	7970.021875	0.00012547017005	0.95
ATtract	Karnani-Ori	255.699102907635	142	0.00010025815601	0.00005567738796	0.99999999999999	0.55534023539884	17002.7133333333	0.00005881414221	0.94666666666667
ATtract	Aladjem-K562	26900.9212775292	715	0.01054769739791	0.00028034741121	1	0.02657901536618	40.5012942465579	0.02469056899545	0.01135443299296
ATtract	Aladjem-MCF7	51239.8478751253	1936	0.02009085133279	0.00075909452883	1	0.03778309421835	27.0758214342587	0.03693332083859	0.02055310791443
ATtract	Bubble	879983.364009645	614008	0.34503644477514	0.24074902554768	1	0.69774955426688	20.6889287278745	0.04833503044808	4.98083943086133
ATtract	ORC	13676.7227073403	3637	0.00536256476215	0.00142604690153	1	0.26592628057363	187.488568698081	0.00533365851019	0.26736749246489
ATtract	G4	19547.7706956818	0	0.00766456910433	0	1	0	7.10584062610576	0.14072930320533	0
ATtract	CpG	26036.8014356894	621	0.01020888094946	0.00024349054876	1	0.02385085593304	89.4565766397755	0.01117860796335	0.02178183093651
ATtract	ATtract	114058.314023492	2550407	0.04472161267730	1	0	22.3605532120587	1	1	1
ATtract	IR	7961.56306654898	54882	0.00312168334957	0.02151891835303	0	6.89337000049529	17.4965835654405	0.05715401502583	0.37650755320170
ATtract	Cruciform	8424.29040711226	64689	0.00330311609367	0.02536418697094	0	7.67886633459192	15.0817061488061	0.06630549555423	0.38253521459912
ATtract	Bends	17129.2126475065	57558	0.00671626632435	0.02256816265011	0	3.36022449977458	46.6159821608086	0.02145187023090	1.05203706750014
ATtract	Z-DNA	8699.50099056614	4556	0.00341102458963	0.00178638154616	1	0.52370819946346	13.6132064393535	0.07345807943594	0.02431838076734
IR	LexoG0	12320.1526936967	3222	0.08452007116677	0.02210391998134	1	0.26152273272136	0.74048899929388	1	0.02210391998134
IR	NS-gDNA	10417.2253892074	2732	0.07146539926462	0.01874236790472	1	0.26225793317579	0.89924613505410	1	0.01874236790472
IR	NS-lexo	5685.04988948681	1396	0.03900120665646	0.00957699326317	1	0.24555633233431	2.18111355508671	0.45848140169861	0.02088850982328
IR	Cadoret	22.3416925567884	9	0.00015327094491	0.00006174279324	0.99878288280470	0.40283429633290	516.900709219858	0.00193460752164	0.03191489361702
IR	Karnani-BrIP	71.0482116856621	52	0.00048741278271	0.00035673613874	0.98893769293631	0.73189738019111	178.853987730061	0.00559115294375	0.06380680981595
IR	Karnani-Lexo	23.3495712368899	25	0.00016018530546	0.00017150775901	0.31824555189122	1.07068347193042	455.51875	0.00219529931534	0.078125
IR	Karnani-Ori	11.4527541107104	15	0.00007856944768	0.00010290465540	0.11873980002590	1.30972863426556	971.773333333333	0.00102904655406	0.1
IR	Aladjem-K562	1225.68974143543	264	0.00840861203185	0.00181112193515	1	0.21538892843373	2.31481157993362	0.43200060370731	0.00419240602817
IR	Aladjem-MCF7	2323.2981952393	429	0.01593854667919	0.00294307314462	1	0.18465128620986	1.5474919050905	0.64620693440171	0.00455438186740
IR	Bubble	39241.2387641607	35800	0.26920707685029	0.24559911090377	1	0.91230555220637	1.18245534338141	0.84569789937296	0.29040998101789
IR	ORC	615.039188916379	298	0.0042193597198	0.00204437248741	1	0.48452197090894	10.7157244725428	0.09332080183307	0.02190693229434
IR	G4	1028.67742501115	106	0.00705704639635	0.00072719289820	1	0.10304493655904	0.40612732191565	1	0.00072719289820
IR	CpG	1172.02107276964	222	0.00804042830817	0.00152298890001	1	0.18941638948127	5.11280252542967	0.19558744837616	0.00778674149421
IR	ATtract	6203.23344546811	62347	0.04255610667417	0.42771977004239	0	10.050726052483	0.05715401502583	1	0.42771977004239
IR	IR	418.783949336473	145766	0.00287298786641	1	0	348.069691378941	1	1	1
IR	Cruciform	449.676620945252	120921	0.00308492118151	0.82955558909485	0	268.906575008982	0.86198005984412	1	0.82955558909485
IR	Bends	786.901218807041	10459	0.00539838658402	0.07175198605984	0	13.2913760330122	2.66429054486301	0.37533444013007	0.19116813803439
IR	Z-DNA	469.972035740251	2567	0.00322415402590	0.01761041669525	0	5.4620271096699	0.77804940538463	1	0.01761041669525
Cruciform	LexoG0	14274.6143735832	2566	0.08441222885990	0.01517391458611	1	0.17975967215959	0.85905583410803	1	0.01517391458611
Cruciform	NS-gDNA	12070.8888439437	2261	0.07138060650682	0.01337031211193	1	0.18731015000062	1.0432331058989	0.95855853724882	0.01394835223136
Cruciform	NS-lexo	6592.89085920838	1112	0.03898673529743	0.00657575721736	1	0.16866652637618	2.53035268064222	0.39520182607358	0.01663898402216
Cruciform	Cadoret	25.9046056991699	11	0.00015318560961	0.00006504795808	0.99917210535176	0.42463491348770	599.666666666667	0.00166759310728	0.03900709219858
Cruciform	Karnani-BrIP	82.3988677193914	56	0.00048726164488	0.00033115324116	0.99868331011506	0.67962098934062	207.492024539877	0.00481946234905	0.06871165644171
Cruciform	Karnani-Lexo	27.0669298008214	28	0.00016005895592	0.00016557662058	0.38019107917546	1.03447270178202	528.45625	0.00189230423521	0.0875
Cruciform	Karnani-Ori	13.2778146485353	21	0.00007851770279	0.00012418246543	0.01738365276956	1.5815855662902	1127.37333333333	0.00088701761025	0.14
Cruciform	Aladjem-K562	1409.34550095311	204	0.00833409518853	0.00120634394994	1	0.14474804074802	2.68545838560607	0.37237590623632	0.00323958647631
Cruciform	Aladjem-MCF7	2677.67516136821	332	0.01583430015119	0.00196326564403	1	0.12398815389928	1.79527575773661	0.55701749198727	0.00352460321673
Cruciform	Bubble	45591.6813891187	40922	0.26960416182228	0.24199023097938	1	0.89757602161535	1.37178967178805	0.72897472591155	0.33195969953112
Cruciform	ORC	711.669863660878	278	0.00420842467837	0.00164393930434	1	0.39063056368573	12.4315224582813	0.08044067034877	0.02043666838197
Cruciform	G4	1106.71812446268	77	0.00654452310658	0.00045533570659	1	0.06957507815044	0.47115628404338	1	0.00045533570659
Cruciform	CpG	1355.51741301328	188	0.00801578544234	0.00111172873818	1	0.13869242710949	5.93146264468607	0.16859248045604	0.00659417748158
Cruciform	ATtract	6577.8669840715	91664	0.03889789235196	0.54205054817688	0	13.9352164192385	0.06630549555423	1	0.54205054817688
Cruciform	IR	450.64263809673	134421	0.00266485303949	0.79489196125507	0	298.287353739365	1.16011964381269	0.86198005984412	0.9221697789608
Cruciform	Cruciform	480.756000498783	169106	0.00284292692452	1	0	351.750159799468	1	1	1
Cruciform	Bends	901.258369541999	30090	0.00532954696783	0.17793573261741	0	33.3866525037555	3.09089579791998	0.32353080316487	0.54998080824697
Cruciform	Z-DNA	499.816338421485	3055	0.00295563929382	0.01806559199555	0	6.11224516919209	0.90263039904349	1	0.01806559199555
Bends	LexoG0	5429.18126420099	5093	0.09923381521450	0.09308914112335	0.99999936194505	0.93807882849340	0.27793102397244	1	0.09308914112335
Bends	NS-gDNA	4571.98088614263	2818	0.08356602668828	0.05150700955932	1	0.61636303173120	0.33751804463966	1	0.05150700955932
Bends	NS-lexo	2402.41751987098	661	0.04391105115737	0.01208166547860	1	0.27513951864432	0.81864703505858	1	0.01208166547860

OverlapAnalyses - OtherNonB

Bends	Cadoret	9.52421247983302	7	0.00017408222258	0.00012794502019	0.73386263145256	0.73496890318460	194.010638297872	0.00515435652793	0.02482269503546
Bends	Karnani-BrIP	29.9373237800502	26	0.00054719021366	0.00047522436073	0.72901546967403	0.86848110375604	67.1300613496932	0.01489645592294	0.03190184049079
Bends	Karnani-Lexo	10.06216731306	10	0.00018391488572	0.00018277860028	0.42473821999001	0.99382167766388	170.971875	0.00584891520900	0.03125
Bends	Karnani-Ori	4.90559646541392	7	0.00008966380554	0.00012794502019	0.12369014221448	1.42694166741034	364.74	0.00274167900422	0.04666666666667
Bends	Aladjem-K562	726.010243157576	234	0.01326991360343	0.00427701924658	1	0.32230950211154	0.86882850836099	1	0.00427701924658
Bends	Aladjem-MCF7	1268.33451504289	546	0.02318244073482	0.00997971157536	1	0.43048580128053	0.58082700780296	1	0.00997971157536
Bends	Bubble	15134.5703869119	14112	0.27662755911812	0.25793716071722	1	0.93243479261253	0.44381621428687	1	0.25793716071722
Bends	ORC	287.209207623257	146	0.00524956969573	0.00266856756411	1	0.50834024858810	4.02198044548997	0.24863372996289	0.01073292656031
Bends	G4	1920.57780219121	584	0.03510405224161	0.01067427025643	1	0.30407515870156	0.15243357099273	1	0.01067427025643
Bends	CpG	558.394864178782	113	0.01020626316789	0.00206539818318	1	0.20236575808444	1.91901087337776	0.52110178940249	0.00396352157137
Bends	ATtract	13235.4996057116	27540	0.24191660919580	0.50337226517519	0	2.08076769449002	0.02145187023090	1	0.50337226517519
Bends	IR	780.375041411563	6664	0.01426358577638	0.12180365922757	0	8.53948376917076	0.37533444013007	1	0.12180365922757
Bends	Cruciform	891.867815606285	9271	0.01630143509726	0.16945404032095	0	10.3950381858971	0.32353080316487	1	0.16945404032095
Bends	Bends	527.29596425787	54711	0.00963784182811	1	0	103.757668763882	1	1	1
Bends	Z-DNA	977.575073470258	2334	0.01786798035989	0.04266052530569	0	2.387540418471	0.29202873796357	1	0.04266052530569
Z-DNA	LexoG0	15802.8136654015	53391	0.08435005265816	0.28498302623993	0	3.3785755581548	0.95172490868728	1	0.28498302623993
Z-DNA	NS-gDNA	13364.0112822639	26886	0.07133255376232	0.14350833742554	0	2.0118211091068	1.15576996631667	0.86522407498345	0.16586262631247
Z-DNA	NS-lexo	7303.38667179993	3713	0.03898299779981	0.01981873305292	1	0.50839427882639	2.80330984124134	0.35672118197151	0.05555804940820
Z-DNA	Cadoret	28.6924952792375	77	0.00015315079573	0.00041099985054	0.00000000000002	2.6836285673529	664.354609929078	0.00150522023186	0.27304964539007
Z-DNA	Karnani-BrIP	91.2826802026386	77	0.00048723594702	0.00041099985054	0.92837407300447	0.84353351401457	229.874846625767	0.00435019322330	0.09447852760736
Z-DNA	Karnani-Lexo	29.9749856923903	44	0.00015999629402	0.00023485705745	0.00617856139915	1.46789060890762	585.4625	0.00170805132694	0.1375
Z-DNA	Karnani-Ori	14.7057313398379	23	0.00007849419977	0.00012276618912	0.01584949485155	1.56401606071049	1248.98666666667	0.00080064905950	0.15333333333333
Z-DNA	Aladjem-K562	1551.76746105421	2108	0.00828280772174	0.01125178811623	0	1.35845096182639	2.97514729002239	0.33611781284027	0.03347572692191
Z-DNA	Aladjem-MCF7	2953.21261036099	4656	0.01576324599334	0.02485214680701	0	1.57658814799347	1.98893784171134	0.50278092106667	0.04942937523223
Z-DNA	Bubble	50570.7993279259	60939	0.26992975280187	0.32527168691419	0	1.20502346828338	1.51976896993689	0.65799474774217	0.49433781657121
Z-DNA	ORC	787.123061856795	2419	0.00420139559459	0.01291180049960	0	3.07321703202758	13.772550172756	0.07260819437624	0.17782842020142
Z-DNA	G4	1158.43166866555	551	0.00618331484011	0.00294105087857	1	0.47564307408370	0.52198140517166	1	0.00294105087857
Z-DNA	CpG	1498.72787678854	10462	0.00799970043335	0.05584260307022	0	6.9805867776463	6.57130831287268	0.15217669709077	0.36695896176780
Z-DNA	ATtract	6804.18168572545	4429	0.03631841111581	0.02364049789696	1	0.65092324170173	0.07345807943594	1	0.02364049789696
Z-DNA	IR	471.773770555525	2607	0.00251816817129	0.01391528065418	0	5.52595367252019	1.28526542540785	0.77804940538463	0.01788482910966
Z-DNA	Cruciform	500.656951957899	2531	0.00267233678479	0.01350961846403	0	5.05535774566222	1.10787316830863	0.90263039904349	0.01496694381039
Z-DNA	Bends	989.52948738137	3591	0.00528177235615	0.01916753848453	0	3.62899746373704	3.42432052055345	0.29202873796357	0.06563579536107
Z-DNA	Z-DNA	518.264676599045	187348	0.00276632083928	1	0	361.490968725506	1	1	1

OverlapAnalyses_ENCODE - Notes/Tables

These analyses were done to get expected proportions, observed proportions, and p-values for ovelraps with Cadoret, Karnani, and Mesner bubble-chip results.

Those analyses were done on the ENCODE regions (1%) of the human genome.

For numPeaks, Proportions, and pValues tables, read as:

this is the numPeaks/proportionPeaks/p-value in/for the ROW set that overlapped peaks in the COLUMN set

obsNumPeaksOver

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	2534	993	106	215	39	63	26	544	793	136	1202	520	548	428	1129
NSgDNAPoolencod	1410	2371	514	152	34	24	10	352	559	90	1126	552	642	312	935
NSlexoPoolencode	109	538	572	6	21	3	3	2	3	3	126	100	60	0	217
Cadoret	204	119	6	282	18	16	6	57	70	32	131	68	84	99	207
KarnaniBrIP	38	34	21	18	815	104	104	16	8	3	231	246	115	2	67
KarnaniLexo	54	21	3	16	107	320	150	13	12	4	86	86	74	13	50
KarnaniOri	23	10	3	6	107	150	150	3	4	2	38	45	37	4	16
AladjemK562encod	784	413	3	63	16	14	3	978	556	28	493	194	240	66	253
AladjemMCF7enco	1234	704	3	85	8	13	4	529	1564	44	854	297	404	130	492
ORCencode	141	81	3	30	3	4	2	26	39	229	127	64	90	84	98
bubbleSeqencode	656	512	113	132	213	85	38	247	369	122	1362	631	388	220	676
bubbleChipGM0699	306	288	89	65	235	86	45	96	136	62	631	988	280	112	392
bubbleChipHela	318	285	49	91	111	70	36	123	179	88	388	280	656	159	371
CpGencode	490	289	0	99	2	13	4	61	120	84	299	139	199	507	292
G4encode	3561	1945	225	455	77	61	21	355	789	203	2622	1212	1450	710	5757

expNumPeaksTable

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	632.834737649822	520.812015534267	144.963448365987	70.9689447354048	215.666068550481	77.2677920633165	37.0467188049739	151.638805912136	256.518702975837	45.9035208100872	929.507032640055	636.56803452655	465.159853637451	96.9722083387784	732.75807892453
NSgDNAPoolencod	520.686884724842	420.481508513386	119.511871545997	58.4533529874772	178.813422267865	63.276415476238	30.434849708913	114.332837379089	195.954636501071	36.4971465155189	831.185192479469	567.678525237052	416.676121823175	76.4473503487603	523.475613342217
NSlexoPoolencode	144.545326319107	119.195798279908	33.1039578045166	16.2081572392782	49.2197347760584	17.6574999528901	8.4631927507719	34.9441097749725	59.037746038861	10.5223769252738	210.345277725768	144.099808526741	105.243304416554	22.2480775494096	169.716532573909
Cadoret	70.734091160783	58.2739175122904	16.2012501975434	7.93197600704383	24.0953567850796	8.63874860283551	4.14119172870609	17.0283778517886	28.7866214729236	5.14044443690414	103.390154517342	70.8180049232556	51.7348082616756	10.8642648774055	82.5057425514284
KarnaniBrIP	215.120865315995	178.40424600332	49.2373239833932	24.1142437118257	73.0799097613664	26.3164374704038	12.6012838319674	53.3174372586343	89.7622915160951	15.8198694650175	304.701062717435	208.936697993565	152.361519319814	33.530756330956	262.566767858752
KarnaniLexo	77.016394646078	63.0856972398758	17.6509609060664	8.63923101597148	26.2972941623151	9.39249425438918	4.50688311016864	18.0682831595188	30.659276561368	5.53941212478519	115.579868766415	79.0967261206964	57.8671223878228	11.6779308760569	86.2366256822224
KarnaniOri	36.916961382764	30.3355390349518	8.45794560786756	4.14038861788379	12.5889722843769	4.50575746522475	2.1608826003062	8.78562567674291	14.8768594026912	2.67046565388241	54.6077157005958	37.3888172921362	27.3318958925487	5.63763660239699	42.2866033853473
AladjemK562encod	151.291729745798	114.0985618426	34.9649789296421	17.0458114631894	53.3302211100474	18.0857714607342	8.79632627047645	22.7322285341843	41.7518713791486	9.32636602286615	308.263077388851	209.098193898643	155.199534075836	18.8550759483667	72.254788587011
AladjemMCF7enco	256.152374961437	195.721822850409	59.1239685769906	28.8409537722534	89.8612732622686	30.715428098016	14.9078293189071	41.7878921435115	75.4677184724279	16.1926743888019	500.919267840401	340.13336456272	252.030898623386	32.9799431045178	147.456308458903
ORCencode	45.7480519029478	36.3822769452919	10.5170737669443	5.14004412071149	15.8062478961514	5.53867144295965	2.6707756389021	9.31610192245014	16.1609110265713	3.11779057823111	77.670179548454	52.9468063577638	38.9826586101508	6.4832759394519	40.441110838761
bubbleSeqencode	927.903674588232	829.95084001254	210.589851611298	103.554498257049	304.946370609556	115.757124438557	54.7051264434224	308.437299374603	500.770342249482	77.7996988352805	814.970303661436	570.954190342097	401.89863934319	169.747729540341	1362
bubbleChipGM0699	635.120377949462	566.523647851257	144.187988094568	70.8915510185906	208.989868654846	79.1744485205868	37.4349062594988	209.101269722869	339.845171879762	53.0059206884899	570.640078652772	399.278820839615	281.640886229092	115.542102956018	988
bubbleChipHela	463.875463387895	415.62546248568	105.256302309284	51.7632563857141	152.325901149529	57.8957095042587	27.3522297353381	155.126058081558	251.694434453243	39.0071322472797	401.481462278772	281.503408302762	197.880952199555	85.1585011487542	656
CpGencode	96.6832700539955	76.2378848977085	22.2459518105971	10.8678581231454	33.515575409897	11.6811409243238	5.64059508724985	18.842021658585	32.9287007655259	6.48592531398908	169.534388629657	115.460409099182	85.1398494525486	13.4306806279434	79.0234611508251
G4encode	736.256640655301	526.10135520678	171.020124107935	83.1749132580394	264.489023391375	86.9312015788357	42.6378417762534	72.7663199726563	148.372226977625	40.7722875057639	1741.10235063172	1176.9428342737	878.467163501526	79.6380490021375	63.4092272181486

obsProportionTable

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	1	0.391870560378848	0.041831097079716	0.084846093133386	0.015390686661405	0.024861878453039	0.01026045777427	0.214680347277032	0.312943962115233	0.053670086819258	0.474348855564325	0.205209155485399	0.216258879242305	0.168902920284136	0.445540647198106
NSgDNAPoolencod	0.594685786587938	1	0.21678616617461	0.064107971320118	0.014339940953184	0.010122311261071	0.004217629692113	0.148460565162379	0.235765499789119	0.037958667229017	0.474905103331927	0.232813159004639	0.270771826233657	0.131590046393927	0.394348376212569
NSlexoPoolencode	0.190559440559441	0.940559440559441	1	0.01048951048951	0.036713286713287	0.005244755244755	0.005244755244755	0.003496503496504	0.005244755244755	0.005244755244755	0.22027972027972	0.174825174825175	0.104895104895105	0	0.379370629370629
Cadoret	0.723404255319149	0.421985815602837	0.021276595744681	1	0.063829787234042	0.056737588652482	0.021276595744681	0.202127659574468	0.24822695035461	0.113475177304965	0.464539007092199	0.24113475177305	0.297872340425532	0.351063829787234	0.734042553191489
KarnaniBrIP	0.046625766871166	0.041717791411043	0.025766871165644	0.022085889570552	1	0.12760736196319	0.12760736196319	0.019631901840491	0.009815950920245	0.003680981595092	0.283435582822086	0.301840490797546	0.141104294478528	0.002453987730061	0.082208588957055
KarnaniLexo	0.16875	0.065625	0.009375	0.05	0.334375	1	0.46875	0.040625	0.0375	0.0125	0.26875	0.26875	0.23125	0.040625	0.15625
KarnaniOri	0.153333333333333	0.066666666666667	0.02	0.04	0.713333333333333	1	1	0.02	0.026666666666667	0.013333333333333	0.253333333333333	0.3	0.246666666666667	0.026666666666667	0.106666666666667
AladjemK562encod	0.801635991820041	0.422290388548057	0.003067484662577	0.06441717791411	0.016359918200409	0.014314928425358	0.003067484662577	1	0.568507157464213	0.028629856850716	0.504089979550102	0.198364008179959	0.245398773006135	0.067484662576687	0.258691206543967
AladjemMCF7enco	0.789002557544757	0.450127877237852	0.001918158567775	0.054347826086956	0.005115089514066	0.008312020460358	0.002557544757033	0.338235294117647	1	0.028132992327366	0.546035805626598	0.189897698209719	0.258312020460358	0.083120204603581	0.314578005115089
ORCencode	0.615720524017467	0.353711790393013	0.013100436681223	0.131004366812227	0.013100436681223	0.017467248908297	0.008733624454148	0.11353711790393	0.170305676855895	1	0.554585152838428	0.279475982532751	0.393013100436681	0.366812227074236	0.427947598253275
bubbleSeqencode	0.481644640234949	0.375917767988253	0.082966226138032	0.096916299559471	0.156387665198238	0.062408223201175	0.027900146842878	0.181350954478708	0.270925110132159	0.089574155653451	1	0.463289280469897	0.284875183553598	0.161527165932452	0.49632892804699
bubbleChipGM0699	0.309716599190283	0.291497975708502	0.090080971659919	0.06578947368421	0.237854251012146	0.087044534412956	0.045546558704453	0.097165991902834	0.137651821862348	0.062753036437247	0.638663967611336	1	0.283400809716599	0.11336032388664	0.396761133603239
bubbleChipHela	0.484756097560976	0.434451219512195	0.07469512195122	0.138719512195122	0.169207317073171	0.106707317073171	0.054878048780488	0.1875	0.272865853658537	0.134146341463415	0.591463414634146	0.426829268292683	1	0.242378048780488	0.565548780487805
CpGencode	0.96646942800789	0.570019723865878	0	0.195266272189349	0.003944773175542	0.025641025641026	0.007889546351085	0.120315581854043	0.236686390532544	0.165680473372781	0.58974358974359	0.274161735700197	0.392504930966469	1	0.575936883629191
G4encode	0.618551328817092	0.337849574431127	0.039082855653986	0.079034219211395	0.013375021712698	0.010595796421747	0.003647733194372	0.061664061142956	0.137050547159979	0.03526142087893	0.455445544554455	0.210526315789474	0.251867291992357	0.123328122285913	1

expProportionTable

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	0.249737465528738	0.205529603604683	0.057207359260453	0.02800668695162	0.085108945757885	0.03049241991449	0.014619857460526	0.059841675577007	0.101230743084387	0.018115043729316	0.366814140741932	0.25121074764268	0.183567424482025	0.038268432651452	0.289170512598473
NSgDNAPoolencod	0.219606446530933	0.177343529529054	0.05040568179924	0.024653459716355	0.07541687991053	0.026687648872306	0.012836292580731	0.048221356971358	0.082646409321413	0.015393144882125	0.35056313474461	0.239425780361473	0.175738558339593	0.03224266147143	0.220782628992921
NSlexoPoolencode	0.252701619438998	0.208384262727112	0.057874052105798	0.028335939229507	0.086048487370731	0.030869755162395	0.014795791522328	0.061091101005197	0.103212842725282	0.018395763855374	0.367736499520573	0.251922742179617	0.183991790938032	0.038895240470996	0.29670722477956
Cadoret	0.250830110499231	0.206645097561314	0.05745124183526	0.028127574493063	0.085444527606665	0.030633860293743	0.014685077052149	0.060384318623364	0.10208021798909	0.018228526372	0.366631753607594	0.251127677032821	0.183456766885374	0.038525761976615	0.292573555146909
KarnaniBrIP	0.263951981982816	0.218900915341497	0.060413894458151	0.029588029094265	0.089668600934192	0.032290107325649	0.015461697953334	0.065420168415502	0.110137781001344	0.019410882779163	0.373866334622619	0.256364046617871	0.186946649472164	0.041142032307922	0.322167813323622
KarnaniLexo	0.240676233268994	0.197142803874612	0.055159252831458	0.026997596924911	0.082179044257235	0.029351544544966	0.014084009719277	0.056463384873496	0.095810239254275	0.017310662889954	0.361187089895046	0.247177269127176	0.180834757461946	0.036493533987678	0.269489455256945
KarnaniOri	0.246113075885093	0.202236926899678	0.05638630405245	0.027602590785892	0.083926481895846	0.030038383101498	0.014405884002041	0.058570837844953	0.099179062684608	0.017803104359216	0.364051438003972	0.249258781947574	0.182212639283658	0.03758424401598	0.281910689235649
AladjemK562encod	0.154695020189978	0.116665196158078	0.035751512197998	0.017429255074836	0.054529878435631	0.018492608855556	0.008994198640569	0.023243587458266	0.042691075029804	0.009536161577573	0.315197420642997	0.213801834252191	0.158690730138892	0.019279218761111	0.073880151929459
AladjemMCF7enco	0.16378029089606	0.125141830467013	0.03780304896227	0.018440507527016	0.057456057073062	0.019639020523028	0.009531860178329	0.026718601114777	0.048253016926105	0.010353372371357	0.320280861790538	0.217476575807366	0.161145075846155	0.021086920143554	0.094281527147636
ORCencode	0.199773152414619	0.158874571813502	0.045926086318534	0.022445607514024	0.06902291657708	0.024186338178863	0.01166277571573	0.040681667783625	0.07057166387149	0.013614806018476	0.33917108973124	0.231208761387615	0.170229950262667	0.028311248643895	0.176598737287166
bubbleSeqencode	0.681280230975207	0.609361850229471	0.154618099567767	0.076031202831901	0.223896013663404	0.084990546577501	0.040165291074466	0.226459103799268	0.367672791666287	0.057121658469369	0.598362924861553	0.419202782923713	0.295079764569155	0.124631225800544	1
bubbleChipGM0699	0.642834390637108	0.573404501873742	0.145939259205028	0.071752582002622	0.211528207140532	0.08013608149857	0.03788958123431	0.211640961257964	0.343972846032148	0.053649717296042	0.577570929810498	0.404128361173699	0.285061625737947	0.116945448336051	1
bubbleChipHela	0.707127230774231	0.63357540013061	0.160451680349518	0.078907403027003	0.23220411760599	0.088255654732102	0.041695472157528	0.236472649514571	0.383680540325066	0.059462091840366	0.612014424205445	0.429121049242015	0.301647792987126	0.129814788336516	1
CpGencode	0.190696785116362	0.150370581652285	0.043877616983426	0.021435617599892	0.066105671419915	0.023039725689002	0.01112543409714	0.037163750805888	0.064948127742655	0.012792752098598	0.334387354299126	0.227732562325803	0.167928697145066	0.026490494335194	0.155864814893146
G4encode	0.1278889422712	0.091384636999614	0.029706465886388	0.014447613906208	0.045942161436751	0.015100087125037	0.007406260513506	0.012639624799836	0.025772490355676	0.007082210787869	0.302432230438026	0.204436830688501	0.152591134879542	0.013833254994292	0.011014282997768

pValueTable

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	0	0	0.99969988970064	0	1	0.947547115023385	0.964879466173656	0	0	0	0	0.999999971992552	0.000013929659706	0	0
NSgDNAPoolencod	0	0	0	0	1	0.999999989743664	0.999984838749566	0	0	0.000000000000013	0	0.766852421944298	0	0	0
NSlexoPoolencode	0.999736625667196	0	0	0.996829866414186	0.99999776210312	0.999980711646383	0.969974603272323	0.99999999999984	1	0.993327855878206	0.999999999999979	0.999993430423403	0.999999869671677	0.999999999860409	0.000009412062269
Cadoret	0	0	0.997131343912327	0	0.886355318538942	0.006779904895342	0.124361755264558	0	0.000000000000612	0	0.000306257629223	0.620929079617021	0.000001097186782	0	0
KarnaniBrIP	1	1	0.999997186948933	0.880007284558251	0	0	0	0.999999999188198	1	0.999902614257214	0.999999967092115	0.001509967198816	0.999684778352482	0.99999999999128	1
KarnaniLexo	0.998845733521701	0.999999999973196	0.999983457901512	0.006887929618831	0	0	0	0.868139158162176	0.999939309433354	0.650663285025052	0.999723845203208	0.168426804825786	0.009373721492961	0.28242068861048	0.999998960196557
KarnaniOri	0.996012214045999	0.999996648249148	0.972257436907254	0.122893609802124	0	0	0	0.978199261065324	0.999393268671766	0.500599538579464	0.997427245413493	0.065206519077935	0.018643430359122	0.668230258655341	0.999999835033027
AladjemK562encod	0	0	0.999999999997034	0	0.999999999058238	0.79993330382978	0.975949650316899	0	0	0.000000158145724	0	0.873053217190029	0.00000000000106	0	0
AladjemMCF7enco	0	0	1	0	1	0.9976016777701	0.999117270541332	0	0	0.000000002398161	0	0.995966399451393	0	0	0
ORCencode	0	0.00000000000017	0.993820926925697	0.000000000000002	0.999925487272811	0.651382781021873	0.500160735880132	0.000000894130875	0.000000107396889	0	0.00000000000813	0.037335612018867	0	0	0
bubbleSeqencode	1	1	0.999999999999998	0.002121514815064	0.999999999693212	0.998882836062502	0.990240871531823	0.999973372555979	0.999999999999977	0.000000608669833	0	0.000462907195244	0.786558615890505	0.000029674704035	1
bubbleChipGM0699	1	1	0.999999915599167	0.743692498516716	0.020561116096806	0.194097548573366	0.092403930656193	1	1	0.092462588126212	0.000039204841305	0	0.530040914482181	0.613938389422315	1
bubbleChipHela	1	1	0.999999999956713	0.000000073256944	0.999952373774773	0.044481707492432	0.041707915051084	0.998527572785374	0.999999998490184	0.00000000000063	0.850807096048807	0.530803063899566	0	0.000000000000002	1
CpGencode	0	0	0.999999999868072	0	0.999999999999407	0.283710563924271	0.665340474719769	0	0	0	0	0.006237442515572	0	0	0
G4encode	0	0	0.00002583404194	0	1	0.998019290442115	0.999818188511245	0	0	0	0	0.122836726237539	0	0	0

OverlapAnalyses_ENCODE - Notes/Tables

setSizeRatioTable

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	1	1.06874736398144	4.43006993006993	8.98581560283688	3.10920245398773	7.91875	16.8933333333333	2.59100204498978	1.62020460358056	11.0655021834061	1.86049926578561	2.56477732793522	3.86280487804878	4.99802761341223	0.44015980545423
NSgDNAPoolencod	0.935674822415154	1	4.14510489510489	8.40780141843972	2.90920245398773	7.409375	15.8066666666667	2.42433537832311	1.51598465473146	10.353711790393	1.74082232011747	2.3997975708502	3.61432926829268	4.67652859960552	0.411846447802675
NSlexoPoolencode	0.225730071033938	0.241248418388865	1	2.02836879432624	0.701840490797546	1.7875	3.81333333333333	0.584867075664622	0.365728900255754	2.49781659388646	0.419970631424376	0.578947368421053	0.871951219512195	1.12820512820513	0.099357304151468
Cadoret	0.111286503551697	0.118937157317588	0.493006993006993	1	0.34601226993865	0.88125	1.88	0.288343558282209	0.180306905370844	1.23144104803493	0.20704845814978	0.285425101214575	0.429878048780488	0.556213017751479	0.048983845752996
KarnaniBrIP	0.32162588792423	0.343736819907212	1.42482517482517	2.89007092198582	1	2.546875	5.43333333333333	0.833333333333333	0.521099744245524	3.5589519650655	0.598384728340675	0.824898785425101	1.24237804878049	1.60749506903353	0.141566788257773
KarnaniLexo	0.126282557221784	0.134964150147617	0.559440559440559	1.13475177304965	0.392638036809816	1	2.13333333333333	0.32719836400818	0.20460358056266	1.39737991266376	0.234948604992658	0.323886639676113	0.48780487804878	0.631163708086785	0.055584505819003
KarnaniOri	0.059194948697711	0.0632644453816696	0.262237762237762	0.531914893617021	0.184049079754601	0.46875	1	0.153374233128834	0.095907928388747	0.655021834061135	0.110132158590308	0.151821862348178	0.228658536585366	0.29585798816568	0.026055237102658
AladjemK562encod	0.385951065509077	0.412484183888655	1.70979020979021	3.46808510638298	1.2	3.05625	6.52	1	0.625319693094629	4.2707423580786	0.718061674008811	0.989878542510121	1.49085365853659	1.92899408284024	0.169880145909328
AladjemMCF7encod	0.617205998421468	0.659637283846478	2.73426573426573	5.54609929078014	1.919018404907 98	4.8875	10.4266666666667	1.59918200408998	1	6.82969432314411	1.14831130690162	1.582995951417	2.38414634146341	3.08481262327416	0.271669272190377
ORCencode	0.090370955011839	0.096583719949388	0.40034965034965	0.812056737588652	0.280981595092025	0.715625	1.52666666666667	0.234151329243354	0.146419437340153	1	0.168135095447871	0.231781376518219	0.349085365853659	0.451676528599605	0.039777661976724
bubbleSeqencode	0.537490134175217	0.574441164065795	2.38111888111888	4.82978723404255	1.67116564417178	4.25625	9.08	1.39263803680982	0.870843989769821	5.94759825327511	1	1.37854251012146	2.07621951219512	2.68639053254438	0.236581552892131
bubbleChipGM0699	0.389897395422257	0.416701813580768	1.72727272727273	3.50354609929078	1.21226993865031	3.0875	6.58666666666667	1.01022494887526	0.631713554987212	4.31441048034934	0.725403817914831	1	1.50609756097561	1.94871794871795	0.171617161716172
bubbleChipHela	0.258879242304657	0.276676507802615	1.14685314685315	2.32624113475177	0.804907975460123	2.05	4.37333333333333	0.670756646216769	0.419437340153453	2.8646288209607	0.481644640234949	0.663967611336032	1	1.29388560157791	0.113948236928956
CpGencode	0.200078926598264	0.213833825390131	0.886363636363636	1.79787234042553	0.622085889570552	1.584375	3.38	0.51840490797546	0.324168797953964	2.21397379912664	0.372246696035242	0.513157894736842	0.772865853658537	1	0.088066701406983
G4encode	2.27190213101815	2.42808941374947	10.0646853146853	20.4148936170213	7.06380368098159	17.990625	38.38	5.88650306748466	3.6809462915601	25.1397379912664	4.22687224669604	5.82692307692308	8.77591463414634	11.3550295857988	1

observed:expected ratio

	LexoPoolencode	NSgDNAPoolencode	NSlexoPoolencode	Cadoret	KarnaniBrIP	KarnaniLexo	KarnaniOri	AladjemK562encode	AladjemMCF7encode	ORCencode	bubbleSeqencode	bubbleChipGM06990	bubbleChipHela	CpGencode	G4encode
LexoPoolencode	4.0042049673357	1.90663803902709	0.731218808567406	3.02949410903021	0.180835122845816	0.815346191701388	0.701816539728483	3.58747219570701	3.09139252148292	2.9627357030556	1.29315858599368	0.816880477491699	1.17808963029539	4.41363569348397	1.54075408033308
NSgDNAPoolencod	2.70796142819137	5.63877352985789	4.30082797090305	2.60036408916632	0.190142325832042	0.379288235899079	0.328570704164552	3.07873055605963	2.85270106378396	2.46594620655431	1.35469208329022	0.972381331087861	1.54076503638105	4.08124020749739	1.78613860162528
NSlexoPoolencode	0.754088719266961	4.51358191952884	17.2789007096294	0.37018397041828	0.426658130027453	0.169899476596564	0.354476152008512	0.057234252435662	0.050814948084656	0.285106684668755	0.599015111545642	0.693963448129377	0.570107526864791	0	1.27860260110782
Cadoret	2.88404073131151	2.04207997471428	0.370341790099004	35.5523011856791	0.747031893345772	1.85212010854769	1.44885829806166	3.34735348816639	2.43168515158478	6.22514266865088	1.26704520959032	0.960207790006097	1.62366504917012	9.11244351248208	2.50891627174885
KarnaniBrIP	0.176644882606722	0.190578423785762	0.426505713573768	0.746446797797454	11.1521757848536	3.95190268884079	8.25312733105578	0.300089442078523	0.089124284428113	0.189634940201866	0.75812009954891	1.17739010122375	0.754783757167 78	0.059646730907566	0.255173190980675
KarnaniLexo	0.701149414331223	0.33288052472734	0.169962418248228	1.85201668648754	4.06885968341694	34.0697573331454	33.2824252001484	0.719492819833928	0.391398667740273	0.722098285863692	0.744074213942956	1.08727635412836	1.27879177236524	1.11321090508025	0.579800051363877
KarnaniOri	0.62301985695763	0.329646359290941	0.354696062033011	1.44913933298046	8.49950238851415	33.2907399383334	69.4160802529246	0.341466858523407	0.26887395328052	0.748933054837194	0.695872360022292	1.20356842658045	1.353729282801	0.709517175743349	0.378370422760041
AladjemK562encod	5.18204135359734	3.61967752555668	0.085800137504359	3.69592261043412	0.300017507277608	0.774089180016218	0.341051469415061	43.0226186811957	13.3167683659247	3.0022411656748	1.59928332700746	0.927793762264817	1.54639639499644	3.50038367284952	3.50149803144647
AladjemMCF7encod	4.81744508590159	3.59694177045384	0.050740842879881	2.94719795576854	0.089026114471475	0.423240072009275	0.268315387467355	12.6591692680566	20.7240927863932	2.71727813105583	1.70486554386663	0.873186905324114	1.60297805628866	3.94178969890915	3.33658156196907
ORCencode	3.08209845304724	2.22635873290173	0.285250447651059	5.8365257759397	0.189798364527 14	0.722194851453869	0.748846129516891	2.79086684714609	2.41323028979477	73.4494489780426	1.63511917621836	1.20876034651739	2.30871888190214	12.9564128975051	2.42327665999895
bubbleSeqencode	0.706969934450479	0.616904008425686	0.536588060323876	1.27469112613864	0.698483472927503	0.734296056612206	0.6946332541485	0.80081105787408	0.736864723941989	1.56812946356388	1.6712265390296	1.10516747345689	0.965417550639378	1.2960408990196	0.496328928046 99
bubbleChipGM0699	0.481798428493107	0.508363598046335	0.617249752743814	0.916893467078382	1.12445642227811	1.08620901827486	1.20208662172306	0.459107685607233	0.400182233714701	1.16968065443798	1.10577581842785	2.47446132485166	0.994173835159157	0.969343617041778	0.396761133603239
bubbleChipHela	0.6855288220625	0.685713522688277	0.465530319087396	1.75800377244263	0.728700760424 44	1.20907059606638	1.3161632652379	0.792903536137893	0.711179809711894	2.25599768376042	0.966420710430184	0.994659360212275	3.31512453678943	1.86710660539057	0.565548780487805
CpGencode	5.06809502539939	3.79076623633727	0	9.1094306604132	0.059673747967615	1.11290498798194	0.709145034899192	3.23744453250884	3.64423731305035	12.9511204544438	1.76365398440288	1.20387586606069	2.33733088887959	37.7493899263121	3.69510517190187
G4encode	4.83662870168555	3.69700625316873	1.31563464342941	5.47039945311902	0.291127393540496	0.7017043235584	0.492520238482028	4.87863066503019	5.31770679777539	4.97887198434261	1.5059424846843	1.02978663424034	1.65060239044152	8.9153364364932	90.7912026777091

OverlapAnalyses_ENCODE - Pairwise-information

fileA	fileB	expNum	obsNum	expProportion	obsProportion	pVal	obsToExpRatio	setSizeAtoBRatio	approxMaxProportio	percentOfMax
LexoPoolencode	LexoPoolencode	632.834737649822	2534	0.24973746552873	1	0	4.0042049673357	1	1	1
LexoPoolencode	NSgDNAPoolencod	520.812015534267	993	0.20552960360468	0.39187056037884	0	1.90663803902709	1.06874736398144	0.93567482241515	0.41881062842682
LexoPoolencode	NSlexoPoolencode	144.963448365987	106	0.05720735926045	0.04183109707971	0.99969988970064	0.73121880856740	4.43006993006993	0.22573007103393	0.18531468531468
LexoPoolencode	Cadoret	70.9689447354048	215	0.02800668695162	0.08484609313338	0	3.02949410903022	8.98581560283688	0.11128650355169	0.76241134751773
LexoPoolencode	KarnaniBrIP	215.666068550481	39	0.08510894575788	0.01539068666140	1	0.18083512284581	3.10920245398773	0.32162588792423	0.04785276073619
LexoPoolencode	KarnaniLexo	77.2677920633165	63	0.03049241991449	0.02486187845303	0.94754711502338	0.81534619170138	7.91875	0.12628255722178	0.196875
LexoPoolencode	KarnaniOri	37.0467188049739	26	0.01461985746052	0.01026045777427	0.96487946617365	0.70181653972845	16.8933333333333	0.05919494869771	0.17333333333333
LexoPoolencode	AladjemK562encod	151.638805912136	544	0.05984167557700	0.21468034727703	0	3.58747219570702	2.59100204498978	0.38595106550907	0.55623721881390
LexoPoolencode	AladjemMCF7encod	256.518702975837	793	0.10123074308438	0.31294396211523	0	3.09139252148291	1.62020460358056	0.61720599842146	0.50703324808184
LexoPoolencode	ORCencode	45.9035208100872	136	0.01811504372931	0.05367008681925	0	2.96273570305558	11.0655021834061	0.09037095501183	0.59388646288209
LexoPoolencode	bubbleSeqencode	929.507032640055	1202	0.36681414074193	0.47434885556432	0	1.29315858599369	1.86049926578561	0.53749013417521	0.88252569750367
LexoPoolencode	bubbleChipGM0699	636.56803452655	520	0.25121074764268	0.20520915548539	0.99999997199255	0.81688047749169	2.56477732793522	0.38989739542225	0.52631578947368
LexoPoolencode	bubbleChipHela	465.159853637451	548	0.18356742448202	0.21625887924230	0.00001392965970	1.17808963029538	3.86280487804878	0.25887924230465	0.83536585365853
LexoPoolencode	CpGencode	96.9722083387784	428	0.03826843265145	0.16890292028413	0	4.41363569348401	4.99802761341223	0.20007892659826	0.84418145956607
LexoPoolencode	G4encode	732.75807892453	1129	0.28917051259847	0.44554064719810	0	1.54075408033308	0.44015980545423	1	0.44554064719810
NSgDNAPoolencod	LexoPoolencode	520.686884724842	1410	0.21960644653093	0.59468578658793	0	2.70796142819137	0.93567482241515	1	0.59468578658793
NSgDNAPoolencod	NSgDNAPoolencod	420.481508513386	2371	0.17734352952905	1	0	5.6387735298579	1	1	1
NSgDNAPoolencod	NSlexoPoolencode	119.511871545997	514	0.05040568179924	0.21678616617461	0	4.30082797090308	4.14510489510489	0.24124841838886	0.89860139860139
NSgDNAPoolencod	Cadoret	58.4533529874772	152	0.02465345971635	0.06410797132011	0	2.60036408916634	8.40780141843972	0.11893715731758	0.53900709219858
NSgDNAPoolencod	KarnaniBrIP	178.813422267865	34	0.07541687991053	0.01433994095318	1	0.19014232583204	2.90920245398773	0.34373681990721	0.04171779141104
NSgDNAPoolencod	KarnaniLexo	63.276415476238	24	0.02668764887230	0.01012231126107	0.99999998974366	0.37928823589908	7.409375	0.13496415014761	0.075
NSgDNAPoolencod	KarnaniOri	30.434849708913	10	0.01283629258073	0.00421762969211	0.99998483874956	0.32857070416455	15.8066666666667	0.06326444538169	0.06666666666667
NSgDNAPoolencod	AladjemK562encod	114.332837379089	352	0.04822135697135	0.14846056516237	0	3.07873055605965	2.42433537832311	0.41248418388865	0.35991820040899
NSgDNAPoolencod	AladjemMCF7encod	195.954636501071	559	0.08264640932141	0.23576549978911	0	2.85270106378394	1.51598465473146	0.65963728384647	0.35741687979539
NSgDNAPoolencod	ORCencode	36.4971465155189	90	0.01539314488212	0.03795866722901	0.00000000000001	2.46594620655429	10.353711790393	0.09658371994938	0.39301310043668
NSgDNAPoolencod	bubbleSeqencode	831.185192479469	1126	0.35056313474461	0.47490510333192	0	1.35469208329023	1.74082232011747	0.57444116406579	0.82672540381791
NSgDNAPoolencod	bubbleChipGM0699	567.678525237052	552	0.23942578036147	0.23281315900463	0.76685242194429	0.97238133108786	2.3997975708502	0.41670181358076	0.55870445344129
NSgDNAPoolencod	bubbleChipHela	416.676121823175	642	0.17573855833959	0.27077182623365	0	1.54076503638105	3.61432926829268	0.27667650780261	0.97865853658536
NSgDNAPoolencod	CpGencode	76.4473503487603	312	0.03224266147143	0.13159004639392	0	4.08124020749739	4.67652859960552	0.21383382539013	0.61538461538461
NSgDNAPoolencod	G4encode	523.475613342217	935	0.22078262899292	0.39434837621256	0	1.78613860162528	0.41184644780267	1	0.39434837621256
NSlexoPoolencode	LexoPoolencode	144.545326319107	109	0.25270161943899	0.19055944055944	0.99973662566719	0.75408871926696	0.22573007103393	1	0.19055944055944
NSlexoPoolencode	NSgDNAPoolencod	119.195798279908	538	0.20838426272711	0.94055944055944	0	4.51358191952884	0.24124841838886	1	0.94055944055944
NSlexoPoolencode	NSlexoPoolencode	33.1039578045166	572	0.05787405210579	1	0	17.2789007096293	1	1	1
NSlexoPoolencode	Cadoret	16.2081572392782	6	0.02833593922950	0.01048951048951	0.99682986641418	0.37018397041829	2.02836879432624	0.49300699300699	0.02127659574468
NSlexoPoolencode	KarnaniBrIP	49.2197347760584	21	0.08604848737073	0.03671328671328	0.99999776210312	0.42665813002744	0.70184049079754	1	0.03671328671328
NSlexoPoolencode	KarnaniLexo	17.6574999528901	3	0.03086975516239	0.00524475524475	0.99998071164638	0.16989947659657	1.7875	0.55944055944055	0.009375
NSlexoPoolencode	KarnaniOri	8.4631927507719	3	0.01479579152232	0.00524475524475	0.96997460327232	0.35447615200851	3.81333333333333	0.26223776223776	0.02
NSlexoPoolencode	AladjemK562encod	34.9441097749725	2	0.06109110100519	0.00349650349650	0.99999999999984	0.05723425243565	0.58486707566462	1	0.00349650349650
NSlexoPoolencode	AladjemMCF7encod	59.037746038861	3	0.10321284272528	0.00524475524475	1	0.05081494808465	0.36572890025575	1	0.00524475524475
NSlexoPoolencode	ORCencode	10.5223769252738	3	0.01839576385537	0.00524475524475	0.99332785878200	0.28510668466877	2.49781659388646	0.40034965034965	0.01310043668122
NSlexoPoolencode	bubbleSeqencode	210.345277725768	126	0.36773649952057	0.22027972027972	0.99999999999997	0.59901511154564	0.41997063142437	1	0.22027972027972
NSlexoPoolencode	bubbleChipGM0699	144.099808526741	100	0.25192274217961	0.17482517482517	0.99999343042340	0.69396344812937	0.57894736842105	1	0.17482517482517
NSlexoPoolencode	bubbleChipHela	105.243304416554	60	0.18399179093803	0.10489510489510	0.99999986967167	0.57010752686478	0.87195121951219	1	0.10489510489510
NSlexoPoolencode	CpGencode	22.2480775494096	0	0.03889524047099	0	0.99999999986040	0	1.12820512820513	0.88636363636363	0
NSlexoPoolencode	G4encode	169.716532573909	217	0.29670722477956	0.37937062937062	0.00000941206226	1.27860260110782	0.09935730415146	1	0.37937062937062
Cadoret	LexoPoolencode	70.734091160783	204	0.25083011049923	0.72340425531914	0	2.88404073131151	0.11128650355169	1	0.72340425531914
Cadoret	NSgDNAPoolencod	58.2739175122904	119	0.20664509756131	0.42198581560283	0	2.04207997471428	0.11893715731758	1	0.42198581560283
Cadoret	NSlexoPoolencode	16.2012501975434	6	0.05745124183526	0.02127659574468	0.99713134391232	0.370341790099	0.49300699300699	1	0.02127659574468
Cadoret	Cadoret	7.93197600704383	282	0.02812757449306	1	0	35.5523011856788	1	1	1
Cadoret	KarnaniBrIP	24.0953567850796	18	0.08544452760666	0.06382978723404	0.88635531853894	0.74703189334577	0.34601226993865	1	0.06382978723404
Cadoret	KarnaniLexo	8.63874860283551	16	0.03063386029374	0.05673758865248	0.00677990489534	1.8521201085477	0.88125	1	0.05673758865248
Cadoret	KarnaniOri	4.14119172870609	6	0.01468507705214	0.02127659574468	0.12436175526455	1.44885829806163	1.88	0.53191489361702	0.04
Cadoret	AladjemK562encod	17.0283778517886	57	0.06038431862336	0.20212765957446	0	3.34735348816641	0.28834355828220	1	0.20212765957446
Cadoret	AladjemMCF7encod	28.7866214729236	70	0.10208021798909	0.24822695035461	0.00000000000061	2.43168515158478	0.18030690537084	1	0.24822695035461
Cadoret	ORCencode	5.14044443690414	32	0.018228526372	0.11347517730496	0	6.22514266865069	1.23144104803493	0.81205673758865	0.13973799126637
Cadoret	bubbleSeqencode	103.390154517342	131	0.36663175360759	0.46453900709219	0.00030625762922	1.26704520959031	0.20704845814978	1	0.46453900709219
Cadoret	bubbleChipGM0699	70.8180049232556	68	0.25112767703282	0.24113475177305	0.62092907961702	0.96020779000609	0.28542510121457	1	0.24113475177305

Cadoret	bubbleChipHela	51.7348082616756	84	0.18345676688537	0.29787234042553	0.00000109718678	1.62366504917012	0.42987804878048	1	0.29787234042553
Cadoret	CpGencode	10.8642648774055	99	0.03852576197661	0.35106382978723	0	9.11244351248199	0.55621301775147	1	0.35106382978723
Cadoret	G4encode	82.5057425514284	207	0.29257355514690	0.73404255319148	0	2.50891627174885	0.04898384575299	1	0.73404255319148
KarnaniBrIP	LexoPoolencode	215.120865315995	38	0.26395198198281	0.04662576687116	1	0.17664488260672	0.32162588792423	1	0.04662576687116
KarnaniBrIP	NSgDNAPoolencod	178.40424600332	34	0.21890091534149	0.04171779141104	1	0.19057842378576	0.34373681990721	1	0.04171779141104
KarnaniBrIP	NSlexoPoolencode	49.2373239833932	21	0.06041389445815	0.02576687116564	0.99999718694893	0.42650571357377	1.42482517482517	0.70184049079754	0.03671328671328
KarnaniBrIP	Cadoret	24.1142437118257	18	0.02958802909426	0.02208588957055	0.88000728455825	0.74644679779746	2.89007092198582	0.34601226993865	0.06382978723404
KarnaniBrIP	KarnaniBrIP	73.0799097613664	815	0.08966860093419	1	0	11.1521757848537	1	1	1
KarnaniBrIP	KarnaniLexo	26.3164374704038	104	0.03229010732564	0.12760736196319	0	3.95190268884082	2.546875	0.39263803680981	0.325
KarnaniBrIP	KarnaniOri	12.6012838319674	104	0.01546169795333	0.12760736196319	0	8.25312733105567	5.43333333333333	0.18404907975460	0.69333333333333
KarnaniBrIP	AladjemK562encod	53.3174372586343	16	0.06542016841550	0.01963190184049	0.99999999918819	0.30008944207851	0.83333333333333	1	0.01963190184049
KarnaniBrIP	AladjemMCF7enco	89.7622915160951	8	0.11013778100134	0.00981595092024	1	0.08912428442811	0.52109974424552	1	0.00981595092024
KarnaniBrIP	ORCencode	15.8198694650175	3	0.01941088277916	0.00368098159509	0.99990261425721	0.18963494020187	3.5589519650655	0.28098159509202	0.01310043668122
KarnaniBrIP	bubbleSeqencode	304.701062717435	231	0.37386633462261	0.28343558282208	0.99999996709211	0.75812009954890	0.59838472834067	1	0.28343558282208
KarnaniBrIP	bubbleChipGM0699	208.936697993565	246	0.25636404661787	0.30184049079754	0.00150996719881	1.17739010122375	0.82489878542510	1	0.30184049079754
KarnaniBrIP	bubbleChipHela	152.361519319814	115	0.18694664947216	0.14110429447852	0.99968477835248	0.75478375716777	1.24237804878049	0.80490797546012	0.17530487804878
KarnaniBrIP	CpGencode	33.530756330956	2	0.04114203230792	0.00245398773006	0.99999999999912	0.05964673090757	1.60749506903353	0.62208588957055	0.00394477317554
KarnaniBrIP	G4encode	262.566767858752	67	0.32216781332362	0.08220858895705	1	0.25517319098067	0.14156678825773	1	0.08220858895705
KarnaniLexo	LexoPoolencode	77.016394646078	54	0.24067623326899	0.16875	0.99884573352170	0.70114941433122	0.12628255722178	1	0.16875
KarnaniLexo	NSgDNAPoolencod	63.0856972398758	21	0.19714280387461	0.065625	0.99999999997319	0.33288052472734	0.13496415014761	1	0.065625
KarnaniLexo	NSlexoPoolencode	17.6509609060664	3	0.05515925283145	0.009375	0.99998345790151	0.16996241824822	0.55944055944055	1	0.009375
KarnaniLexo	Cadoret	8.63923101597148	16	0.02699759692491	0.05	0.00688792961883	1.85201668648755	1.13475177304965	0.88125	0.05673758865248
KarnaniLexo	KarnaniBrIP	26.2972941623151	107	0.08217904425723	0.334375	0	4.06885968341695	0.39263803680981	1	0.334375
KarnaniLexo	KarnaniLexo	9.39249425438918	320	0.02935154454496	1	0	34.0697573331452	1	1	1
KarnaniLexo	KarnaniOri	4.50688311016864	150	0.01408400971927	0.46875	0	33.2824252001484	2.13333333333333	0.46875	1
KarnaniLexo	AladjemK562encod	18.0682831595188	13	0.05646338487349	0.040625	0.86813915816217	0.71949281983392	0.32719836400818	1	0.040625
KarnaniLexo	AladjemMCF7enco	30.659276561368	12	0.09581023925427	0.0375	0.99993930943335	0.39139866774027	0.20460358056266	1	0.0375
KarnaniLexo	ORCencode	5.53941212478519	4	0.01731066288995	0.0125	0.65066328502505	0.72209828586370	1.39737991266376	0.715625	0.01746724890829
KarnaniLexo	bubbleSeqencode	115.579868766415	86	0.36118708989504	0.26875	0.99972384520320	0.74407421394295	0.23494860499265	1	0.26875
KarnaniLexo	bubbleChipGM0699	79.0967261206964	86	0.24717726912717	0.26875	0.16842680482578	1.08727635412836	0.32388663967611	1	0.26875
KarnaniLexo	bubbleChipHela	57.8671223878228	74	0.18083475746194	0.23125	0.00937372149296	1.27879177236524	0.48780487804878	1	0.23125
KarnaniLexo	CpGencode	11.6779308760569	13	0.03649353398767	0.040625	0.28242068861048	1.11321090508026	0.63116370808678	1	0.040625
KarnaniLexo	G4encode	86.2366256822224	50	0.26948945525694	0.15625	0.99999896019655	0.57980005136387	0.05558450581900	1	0.15625
KarnaniOri	LexoPoolencode	36.916961382764	23	0.24611307588509	0.15333333333333	0.99601221404599	0.62301985695763	0.05919494869771	1	0.15333333333333
KarnaniOri	NSgDNAPoolencod	30.3355390349518	10	0.20223692689967	0.06666666666666	0.99999664824914	0.32964635929093	0.06326444538169	1	0.06666666666667
KarnaniOri	NSlexoPoolencode	8.45794560786756	3	0.05638630405245	0.02	0.97225743690725	0.35469606203300	0.26223776223776	1	0.02
KarnaniOri	Cadoret	4.14038861788379	6	0.02760259078589	0.04	0.12289360980212	1.44913933298046	0.53191489361702	1	0.04
KarnaniOri	KarnaniBrIP	12.5889722843769	107	0.08392648189584	0.71333333333333	0	8.49950238851418	0.18404907975460	1	0.71333333333333
KarnaniOri	KarnaniLexo	4.50575746522475	150	0.03003838310149	1	0	33.2907399383331	0.46875	1	1
KarnaniOri	KarnaniOri	2.1608826003062	150	0.01440588400204	1	0	69.4160802529228	1	1	1
KarnaniOri	AladjemK562encod	8.78562567674291	3	0.05857083784495	0.02	0.97819926106532	0.34146685852340	0.15337423312883	1	0.02
KarnaniOri	AladjemMCF7enco	14.8768594026912	4	0.09917906268460	0.02666666666666	0.99939326867176	0.26887395328051	0.09590792838874	1	0.02666666666667
KarnaniOri	ORCencode	2.67046565388241	2	0.01780310435921	0.01333333333333	0.50059953857946	0.74893305483721	0.65502183406113	1	0.01333333333333
KarnaniOri	bubbleSeqencode	54.6077157005958	38	0.36405143800397	0.25333333333333	0.99742724541349	0.69587236002229	0.11013215859030	1	0.25333333333333
KarnaniOri	bubbleChipGM0699	37.3888172921362	45	0.24925878194757	0.3	0.06520651907793	1.20356842658044	0.15182186234817	1	0.3
KarnaniOri	bubbleChipHela	27.3318958925487	37	0.18221263928365	0.24666666666666	0.01864343035912	1.3537297282801	0.22865853658536	1	0.24666666666667
KarnaniOri	CpGencode	5.63763660239699	4	0.03758424401598	0.02666666666666	0.66823025865534	0.70951717574334	0.29585798816568	1	0.02666666666667
KarnaniOri	G4encode	42.2866033853473	16	0.28191068923564	0.10666666666666	0.99999983503302	0.37837042276004	0.02605523710265	1	0.10666666666667
AladjemK562encod	LexoPoolencode	151.291729745798	784	0.15469502018997	0.80163599182004	0	5.18204135359734	0.38595106550907	1	0.80163599182004
AladjemK562encod	NSgDNAPoolencod	114.0985618426	413	0.11666519615807	0.42229038854805	0	3.61967752555669	0.41248418388865	1	0.42229038854805
AladjemK562encod	NSlexoPoolencode	34.9649789296421	3	0.03575151219799	0.00306748466257	0.99999999997034	0.08580013750435	1.70979020979021	0.58486707566462	0.00524475524475
AladjemK562encod	Cadoret	17.0458114631894	63	0.01742925507483	0.06441717791411	0	3.6959226104342	3.46808510638298	0.28834355828220	0.22340425531914
AladjemK562encod	KarnaniBrIP	53.3302211100474	16	0.05452987843563	0.01635991820040	0.99999999905823	0.30001750727760	1.2	0.83333333333333	0.01963190184049
AladjemK562encod	KarnaniLexo	18.0857714607342	14	0.01849260885555	0.01431492842535	0.79993330382978	0.77408918001619	3.05625	0.32719836400818	0.04375
AladjemK562encod	KarnaniOri	8.79632627047645	3	0.00899419864056	0.00306748466257	0.97594965031689	0.34105146941502	6.52	0.15337423312883	0.02
AladjemK562encod	AladjemK562encod	22.7322285341843	978	0.02324358745826	1	0	43.0226186811954	1	1	1
AladjemK562encod	AladjemMCF7enco	41.7518713791486	556	0.04269107502980	0.56850715746421	0	13.3167683659246	0.62531969309462	1	0.56850715746421
AladjemK562encod	ORCencode	9.32636602286615	28	0.00953616157757	0.02862985685071	0.00000015814572	3.00224116567485	4.2707423580786	0.23415132924335	0.12227074235807

OverlapAnalyses_ENCODE - Pairwise-information

AladjemK562encod	bubbleSeqencode	308.263077388851	493	0.31519742064299	0.50408997955010	0	1.59928332700746	0.71806167400881	1	0.50408997955010
AladjemK562encod	bubbleChipGM0699	209.098193898643	194	0.21380183425219	0.19836400817995	0.87305321719002	0.92779376226481	0.98987854251012	1	0.19836400817995
AladjemK562encod	bubbleChipHela	155.199534075836	240	0.15869073013889	0.24539877300613	0.00000000000106	1.54639639499644	1.49085365853659	0.67075664621676	0.36585365853658
AladjemK562encod	CpGencode	18.8550759483667	66	0.01927921876111	0.06748466257668	0	3.5003836728495	1.92899408284024	0.51840490797546	0.13017751479289
AladjemK562encod	G4encode	72.254788587011	253	0.07388015192945	0.25869120654396	0	3.50149803144647	0.16988014590932	1	0.25869120654396
AladjemMCF7encod	LexoPoolencode	256.152374961437	1234	0.16378029089606	0.78900255754475	0	4.8174450859016	0.61720599842146	1	0.78900255754475
AladjemMCF7encod	NSgDNAPoolencod	195.721822850409	704	0.12514183046701	0.45012787723785	0	3.59694177045383	0.65963728384647	1	0.45012787723785
AladjemMCF7encod	NSlexoPoolencode	59.1239685769906	3	0.03780304896227	0.00191815856777	1	0.05074084287987	2.73426573426573	0.36572890025575	0.00524475524475
AladjemMCF7encod	Cadoret	28.8409537722534	85	0.01844050752701	0.05434782608695	0	2.94719795576853	5.54609929078014	0.18030690537084	0.30141843971631
AladjemMCF7encod	KarnaniBrIP	89.8612732622686	8	0.05745605707306	0.00511508951406	1	0.08902611447148	1.91901840490798	0.52109974424552	0.00981595092024
AladjemMCF7encod	KarnaniLexo	30.7154280980166	13	0.01963902052302	0.00831202046035	0.99976016777701	0.42324007200926	4.8875	0.20460358056266	0.040625
AladjemMCF7encod	KarnaniOri	14.9078293189071	4	0.00953186017832	0.00255754475703	0.99911727054133	0.26831538746737	10.4266666666667	0.09590792838874	0.02666666666667
AladjemMCF7encod	AladjemK562encod	41.7878921435115	529	0.02671860111477	0.33823529411764	0	12.6591692680565	1.59918200408998	0.62531969309462	0.54089979550102
AladjemMCF7encod	AladjemMCF7encod	75.4677184724279	1564	0.04825301692610	1	0	20.7240927863933	1	1	1
AladjemMCF7encod	ORCencode	16.1926743888019	44	0.01035337237135	0.02813299232736	0.00000000239816	2.71727813105589	6.82969432314411	0.14641943734015	0.19213973799126
AladjemMCF7encod	bubbleSeqencode	500.919267840401	854	0.32028086179053	0.54603580562659	0	1.70486554386663	1.14831130690162	0.87084398976982	0.62701908957415
AladjemMCF7encod	bubbleChipGM0699	340.13336456272	297	0.21747657580736	0.18989769820971	0.99596639945139	0.87318690532411	1.582995951417	0.63171355498721	0.30060728744939
AladjemMCF7encod	bubbleChipHela	252.030898623386	404	0.16114507584615	0.25831202046035	0	1.60297805628866	2.38414634146341	0.41943734015345	0.61585365853658
AladjemMCF7encod	CpGencode	32.9799431045178	130	0.02108692014355	0.08312020460358	0	3.9417896989092	3.08481262327416	0.32416879795396	0.25641025641025
AladjemMCF7encod	G4encode	147.456308458903	492	0.09428152714763	0.31457800511508	0	3.33658156196906	0.27166927219037	1	0.31457800511508
ORCencode	LexoPoolencode	45.7480519029478	141	0.19977315241461	0.61572052401746	0	3.08209845304723	0.09037095501183	1	0.61572052401746
ORCencode	NSgDNAPoolencod	36.3822769452919	81	0.15887457181350	0.35371179039301	0.00000000000017	2.22635873290173	0.09658371994938	1	0.35371179039301
ORCencode	NSlexoPoolencode	10.5170737669443	3	0.04592608631853	0.01310043668122	0.99382092692569	0.28525044765105	0.40034965034965	1	0.01310043668122
ORCencode	Cadoret	5.14004412071149	30	0.02244560751402	0.13100436681222	0.00000000000000	5.83652577593972	0.81205673758865	1	0.13100436681222
ORCencode	KarnaniBrIP	15.8062478961514	3	0.06902291657708	0.01310043668122	0.99992548727281	0.18979836452713	0.28098159509202	1	0.01310043668122
ORCencode	KarnaniLexo	5.53867144295965	4	0.02418633817886	0.01746724890829	0.65138278102187	0.72219485145386	0.715625	1	0.01746724890829
ORCencode	KarnaniOri	2.6707756389021	2	0.01166277571573	0.00873362445414	0.50016073588013	0.74884612951695	1.52666666666667	0.65502183406113	0.01333333333333
ORCencode	AladjemK562encod	9.31610192245014	26	0.04068166778362	0.11353711790393	0.00000089413087	2.79086684714609	0.23415132924335	1	0.11353711790393
ORCencode	AladjemMCF7encod	16.1609110265713	39	0.07057166387149	0.17030567685589	0.00000010739688	2.41323028979476	0.14641943734015	1	0.17030567685589
ORCencode	ORCencode	3.11779057823111	229	0.01361480601847	1	0	73.4494489780401	1	1	1
ORCencode	bubbleSeqencode	77.670179548454	127	0.33917108973124	0.55458515283842	0.00000000000813	1.63511917621836	0.16813509544787	1	0.55458515283842
ORCencode	bubbleChipGM0699	52.9468063577638	64	0.23120876138761	0.27947598253275	0.03733561201886	1.2087603465174	0.23178137651821	1	0.27947598253275
ORCencode	bubbleChipHela	38.9826586101508	90	0.17022995026266	0.39301310043668	0	2.30871888190213	0.34908536585365	1	0.39301310043668
ORCencode	CpGencode	6.4832759394519	84	0.02831124864389	0.36681222707423	0	12.9564128975052	0.45167652859605	1	0.36681222707423
ORCencode	G4encode	40.441110838761	98	0.17659873728716	0.42794759825327	0	2.42327665999895	0.03977766197672	1	0.42794759825327
bubbleSeqencode	LexoPoolencode	927.903674588232	656	0.68128023097520	0.48164464023494	1	0.70696993445047	0.53749013417521	1	0.48164464023494
bubbleSeqencode	NSgDNAPoolencod	829.95084001254	512	0.60936185022947	0.37591776798825	1	0.61690400842568	0.57444116406579	1	0.37591776798825
bubbleSeqencode	NSlexoPoolencode	210.589851611298	113	0.15461809956776	0.08296622613803	0.99999999999999	0.53658806032387	2.38111888111888	0.41997063142437	0.19755244755244
bubbleSeqencode	Cadoret	103.554498257049	132	0.07603120283190	0.09691629955947	0.00212151481506	1.27469112613865	4.82978723404255	0.20704845814978	0.46808510638297
bubbleSeqencode	KarnaniBrIP	304.946370609556	213	0.22389601366340	0.15638766519823	0.99999999969321	0.69848347292750	1.67116564417178	0.59838472834067	0.26134969325153
bubbleSeqencode	KarnaniLexo	115.757124438557	85	0.08499054657750	0.06240822320117	0.99888283606250	0.73429605661220	4.25625	0.23494860499265	0.265625
bubbleSeqencode	KarnaniOri	54.7051264434224	38	0.04016529107446	0.02790014684287	0.99024087153182	0.69463325414850	9.08	0.11013215859030	0.25333333333333
bubbleSeqencode	AladjemK562encod	308.437299374603	247	0.22645910379926	0.18135095447870	0.99997337255975	0.80081105787407	1.39263803680982	0.71806167400881	0.25255623721881
bubbleSeqencode	AladjemMCF7encod	500.770342249482	369	0.36767279166628	0.27092511013215	0.99999999999997	0.73686472394198	0.87084398976982	1	0.27092511013215
bubbleSeqencode	ORCencode	77.7996988352805	122	0.05712165846936	0.08957415565345	0.00000060866983	1.56812946356388	5.94759825327511	0.16813509544787	0.53275109170305
bubbleSeqencode	bubbleSeqencode	814.970303661436	1362	0.59836292486155	1	0	1.6712265390296	1	1	1
bubbleSeqencode	bubbleChipGM0699	570.954190342097	631	0.41920278292371	0.46328928046989	0.00046290719524	1.10516747345689	1.37854251012146	0.72540381791483	0.63866396761133
bubbleSeqencode	bubbleChipHela	401.89863934319	388	0.29507976456915	0.28487518355359	0.78655861589050	0.96541755063937	2.07621951219512	0.48164464023494	0.59146341463414
bubbleSeqencode	CpGencode	169.747729540341	220	0.12463122580054	0.16152716593245	0.00002967470403	1.29604089901961	2.68639053254438	0.37224669603524	0.43392504930966
bubbleSeqencode	G4encode	1362	676	1	0.49632892804699	1	0.49632892804699	0.23658155289213	1	0.49632892804699
bubbleChipGM0699	LexoPoolencode	635.120377949462	306	0.64283439063710	0.30971659919028	1	0.48179842849310	0.38989739542225	1	0.30971659919028
bubbleChipGM0699	NSgDNAPoolencod	566.523647851257	288	0.57340450187374	0.29149797570850	1	0.50836359804633	0.41670181358076	1	0.29149797570850
bubbleChipGM0699	NSlexoPoolencode	144.187988094568	89	0.14593925920502	0.09008097165991	0.99999991559916	0.61724975274381	1.72727272727273	0.57894736842105	0.15559440559440
bubbleChipGM0699	Cadoret	70.8915510185906	65	0.07175258200262	0.06578947368421	0.74369249851671	0.91689346707838	3.50354609929078	0.28542510121457	0.23049645390070
bubbleChipGM0699	KarnaniBrIP	208.989868654846	235	0.21152820714053	0.23785425101214	0.02056111609680	1.12445642227811	1.21226993865031	0.82489878542510	0.28834355828220
bubbleChipGM0699	KarnaniLexo	79.1744485205868	86	0.08013608149857	0.08704453441295	0.19409754857336	1.08620901827486	3.0875	0.32388663967611	0.26875
bubbleChipGM0699	KarnaniOri	37.4349062594988	45	0.03788958123431	0.04554655870445	0.09240393065619	1.20208662172305	6.58666666666667	0.15182186234817	0.3
bubbleChipGM0699	AladjemK562encod	209.101269722869	96	0.21164096125796	0.09716599190283	1	0.45910768560723	1.01022494887526	0.98987854251012	0.09815950920245

OverlapAnalyses_ENCODE - Pairwise-information

bubbleChipGM0699	AladjemMCF7encod	339.845171879762	136	0.343972846032148	0.137651821862348	1	0.400182233714702	0.631713554987212	1	0.137651821862348
bubbleChipGM0699	ORCencode	53.0059206884899	62	0.053649717296042	0.062753036437247	0.092462588126213	1.16968065443797	4.31441048034934	0.231781376518219	0.270742358078603
bubbleChipGM0699	bubbleSeqencode	570.640078652772	631	0.577570929810496	0.638663967611336	0.000039204841304	1.10577581842785	0.72540381791483	1	0.638663967611336
bubbleChipGM0699	bubbleChipGM0699	399.278820839615	988	0.404128361173699	1	0	2.47446132485166	1	1	1
bubbleChipGM0699	bubbleChipHela	281.640886229092	280	0.285061625737947	0.283400809716599	0.530040914482188	0.994173835159156	1.50609756097561	0.663967611336033	0.426829268292683
bubbleChipGM0699	CpGencode	115.542102956018	112	0.116945448336054	0.113360323886647	0.613938389422315	0.969343617041770	1.94871794871795	0.513157894736842	0.220907297830375
bubbleChipGM0699	G4encode	988	392	1	0.396761133603239	1	0.396761133603239	0.171617161716172	1	0.396761133603239
bubbleChipHela	LexoPoolencode	463.875463387895	318	0.707127230774234	0.484756097560976	1	0.685528822062495	0.258879242304657	1	0.484756097560976
bubbleChipHela	NSgDNAPoolencod	415.62546248568	285	0.633575400130614	0.434451219512195	1	0.685713522688277	0.276676507802615	1	0.434451219512195
bubbleChipHela	NSlexoPoolencode	105.256302309284	49	0.160451680349518	0.074695121951220	0.999999999956713	0.465530319087392	1.14685314685315	0.871951219512195	0.085664335664336
bubbleChipHela	Cadoret	51.7632563857141	91	0.078907403027006	0.138719512195122	0.000000073256944	1.75800377244262	2.32624113475177	0.429878047804878	0.322695034560993
bubbleChipHela	KarnaniBrIP	152.325901149529	111	0.232204117605995	0.169207317073170	0.999952373774775	0.728700760424435	0.804907975460123	1	0.169207317073171
bubbleChipHela	KarnaniLexo	57.8957095042587	70	0.088255654732102	0.106707317073170	0.044481704924332	1.20907059606638	2.05	0.48780487804878	0.21875
bubbleChipHela	KarnaniOri	27.3522297353381	36	0.041695472157528	0.054878048780488	0.041707915051084	1.31616326523791	4.37333333333333	0.228658536585366	0.24
bubbleChipHela	AladjemK562encod	155.126058081558	123	0.236472649514577	0.1875	0.998527572785374	0.792903536137894	0.670756646216769	1	0.1875
bubbleChipHela	AladjemMCF7encod	251.694434453243	179	0.383680540325066	0.272865853658537	0.999999998490184	0.711179809711895	0.419437340153453	1	0.272865853658537
bubbleChipHela	ORCencode	39.0071322472797	88	0.059462091840360	0.134146341463415	0.000000000000632	2.25599768376043	2.8646288209607	0.349085365853659	0.384279475982533
bubbleChipHela	bubbleSeqencode	401.481462278772	388	0.612014424205445	0.591463414634146	0.850807096048801	0.966420710430184	0.481644640234949	1	0.591463414634146
bubbleChipHela	bubbleChipGM0699	281.503408302762	280	0.429121049242015	0.426829268292683	0.530803063899560	0.994659360212270	0.663967611336032	1	0.426829268292683
bubbleChipHela	bubbleChipHela	197.880952199555	656	0.301647792987120	1	0	3.31512453678943	1	1	1
bubbleChipHela	CpGencode	85.1585011487542	159	0.129814788336510	0.242378048780488	0.000000000000003	1.86710660539058	1.29388560157791	0.772865853658537	0.313609467455621
bubbleChipHela	G4encode	656	371	1	0.565548780487805	1	0.565548780487805	0.113948236928956	1	0.565548780487805
CpGencode	LexoPoolencode	96.6832700539955	490	0.190696785116360	0.966469428007890	0	5.06809502539939	0.200078926598264	1	0.966469428007890
CpGencode	NSgDNAPoolencod	76.2378848977085	289	0.150370581652280	0.570019723865877	0	3.79076623633726	0.213833825390131	1	0.570019723865878
CpGencode	NSlexoPoolencode	22.2459518105971	0	0.043877616983420	0	0.999999998680727	0	0.886363636363636	1	0
CpGencode	Cadoret	10.8678581231454	99	0.021435617599893	0.195266272183945	0	9.10943066041311	1.79787234042553	0.556213017751479	0.351063829787234
CpGencode	KarnaniBrIP	33.515575409897	2	0.066105671419915	0.003944773175542	0.999999999994010	0.05967374796762	0.622085889570552	1	0.003944773175542
CpGencode	KarnaniLexo	11.6811409243238	13	0.023039725689000	0.025641025641020	0.283710563924277	1.11290498798195	1.584375	0.631163708086785	0.040625
CpGencode	KarnaniOri	5.64059508724985	4	0.011125434097140	0.007889546351085	0.665340474719760	0.70914503489919	3.38	0.295857988165680	0.026666666666667
CpGencode	AladjemK562encod	18.842021658585	61	0.037163750805880	0.120315581854043	0	3.23744453250888	0.51840490797546	1	0.120315581854043
CpGencode	AladjemMCF7encod	32.9287007655259	120	0.064948127742650	0.236686390532544	0	3.64423731305038	0.324168797953964	1	0.236686390532544
CpGencode	ORCencode	6.48592531398908	84	0.012792752098590	0.165680473372781	0	12.951120454444	2.21397379912664	0.451676528599600	0.366812227074236
CpGencode	bubbleSeqencode	169.534388629657	299	0.334387354299120	0.589743589743590	0	1.76365398440287	0.372246696035242	1	0.589743589743590
CpGencode	bubbleChipGM0699	115.460409099182	139	0.227732562325800	0.274161735700197	0.006237442515573	1.20387586606069	0.513157894736842	1	0.274161735700197
CpGencode	bubbleChipHela	85.1398494525486	199	0.167928697145060	0.392504930966469	0	2.33733088887959	0.772865853658537	1	0.392504930966469
CpGencode	CpGencode	13.4306806279434	507	0.026490494335194	1	0	37.7493899263119	1	1	1
CpGencode	G4encode	79.0234611508251	292	0.155864814893140	0.575936883629191	0	3.69510517190186	0.088066701406983	1	0.575936883629191
G4encode	LexoPoolencode	736.256640655301	3561	0.127888942271200	0.618551328817092	0	4.83662870168553	2.27190213101815	0.440159805454230	1.405288082083660
G4encode	NSgDNAPoolencod	526.10135520678	1945	0.091384636999614	0.337849574431127	0	3.69700625316871	2.42808941374947	0.411846447802675	0.820328975115985
G4encode	NSlexoPoolencode	171.020124107935	225	0.029706465886380	0.039082855653980	0.000025834041940	1.31563464342943	10.0646853146853	0.099357304151463	0.393356643356643
G4encode	Cadoret	83.1749132580394	455	0.014447613906200	0.079034219211395	0	5.47039945311901	20.4148936170213	0.048983847529960	1.613475177304965
G4encode	KarnaniBrIP	264.489023391375	77	0.045942161436750	0.013375021712698	1	0.291127393540483	7.06380368098159	0.141566788257773	0.094478527607362
G4encode	KarnaniLexo	86.9312015788357	61	0.015100087125037	0.010595796421741	0.998019290442115	0.701704323558448	17.990625	0.055584505819003	0.190625
G4encode	KarnaniOri	42.6378417762534	21	0.007406260513500	0.003647733194372	0.999818188511245	0.492520238482044	38.38	0.026055237102653	0.14
G4encode	AladjemK562encod	72.7663199726563	355	0.012639624799830	0.061664061142950	0	4.87863066503019	5.88650306748466	0.169880145909322	0.362985685071575
G4encode	AladjemMCF7encod	148.372226977625	789	0.025772490355670	0.137050547159975	0	5.31770679777546	3.6809462915601	0.271669272190377	0.504475703324808
G4encode	ORCencode	40.7722875057639	203	0.007082210787860	0.035261420878930	0	4.97887198434236	25.1397379912664	0.039777661976724	0.886462882096070
G4encode	bubbleSeqencode	1741.10235063172	2622	0.302432230438020	0.455445544554455	0	1.5059424846843	4.22687224669604	0.236581552892133	1.925110132158590
G4encode	bubbleChipGM0699	1176.9428342737	1212	0.204436830688500	0.210526315789474	0.122836726237539	1.02978663424034	5.82692307692308	0.171617161716171	1.226720647773280
G4encode	bubbleChipHela	878.467163501526	1450	0.152591134879540	0.251867291992357	0	1.65060239044152	8.77591463414634	0.113948236928956	2.210365853658540
G4encode	CpGencode	79.6380490021375	710	0.013833254994290	0.123328122285913	0	8.91533643649336	11.3550295857988	0.088066701406983	1.400394477317554
G4encode	G4encode	63.4092272181486	5757	0.011014282997768	1	0	90.7912026777117	1	1	1

Appendix 2.

File	Starting # reads	Keep 1 read leaves:	Keep 3 reads leaves:
NS-Comb-Align	239,583,014	171,804,663	212,189,231
Input	179,965,523	165182107	176,806,690

Figure 1. Number of redundant reads retained by MACS.
First, MACS has options on how to deal with "redundant reads". Redundant reads are reads that map to the same genomic position on the same strand. It is thought that this is an artifact, usually a consequence of the PCR amplification steps in next-generation sequencing preparations. MACS allows one to keep all occurrences, though this is advised against (since it is likely an artifact), as well as two other options: keep only 1 of the reads at that site (discard all others) or use the binomial distribution to determine how many reads could reasonably pile up at the same site on the same strand. In our case, the binomial option allows 3 reads to be kept at a site. We explored these latter two options. Henceforth, we will use the term "K1" to refer to the option that keeps only 1 read and "K3" to be the binomial option that allows 3 reads to be kept in our data. Note that all but 1 or 3 reads at a redundant site are discarded and this reduces the number of reads in the treatment and control files used to call peaks. The table presented here in Figure 2 shows how many are left in each file for each option.

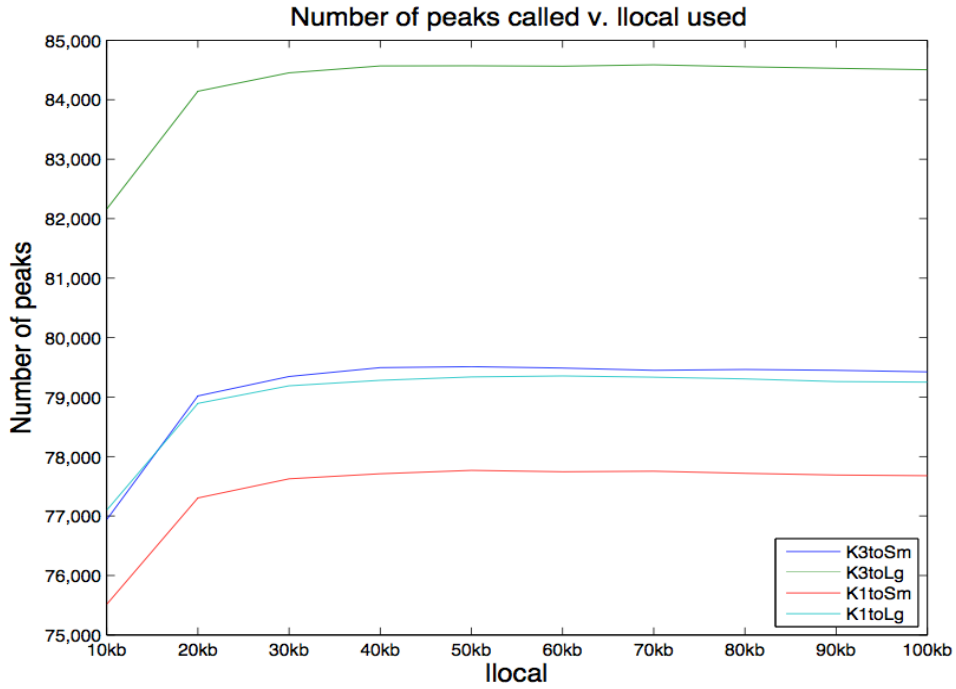


Figure 2. Effect of varying the llocal value on the number of peaks called. Note that all conditions have a slight elevation in number of peaks at llocal = ~50kb.

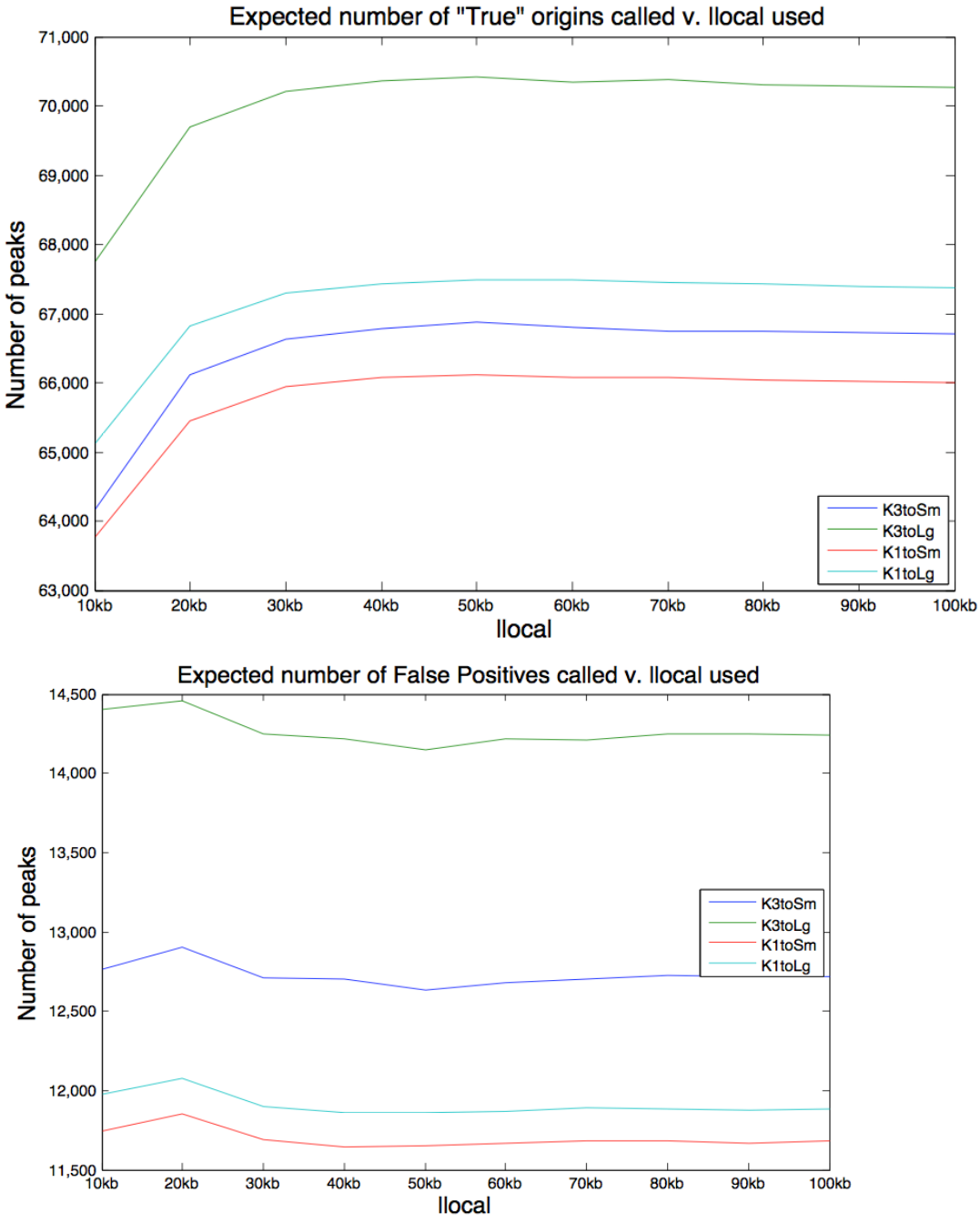


Figure 3. False Discovery Rate (FDR) of the peaks called.
Note that the expected number of true origins based on the data stayed somewhat constant for all 4 conditions after llocal=~30kb with a range from ~66,000 to ~70,000. Moreover, note that all conditions had a slight elevation of true peaks and a slight dip in false peaks around llocal=50kb. As will be seen below, taken together this means that there is also a slight decrease in FDR at this llocal value.

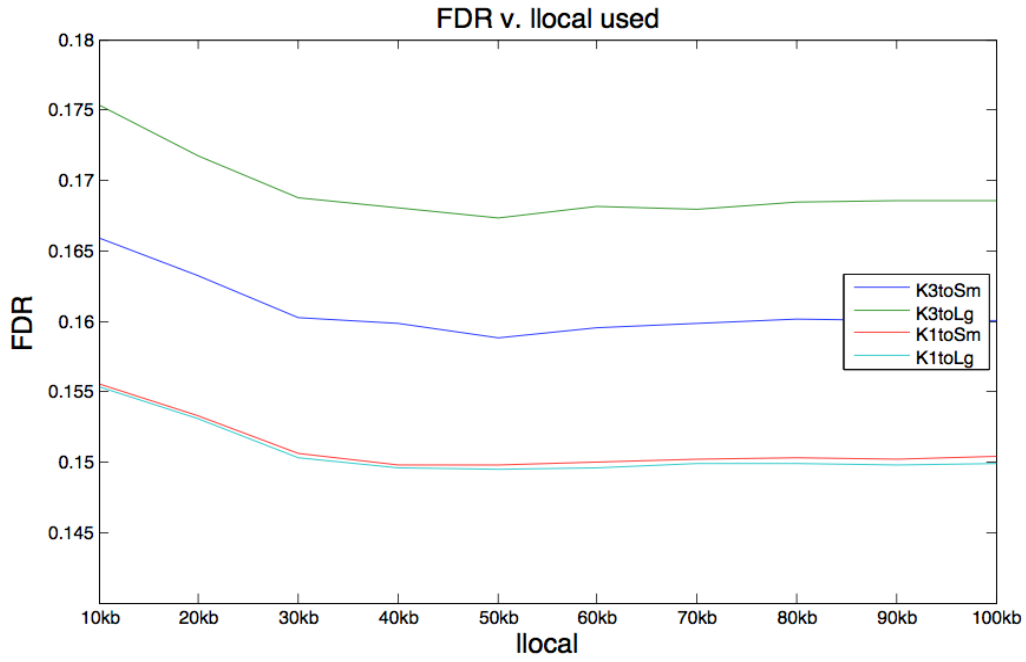


Figure 4. Effect of varying llocal on the False Discovery Rate.

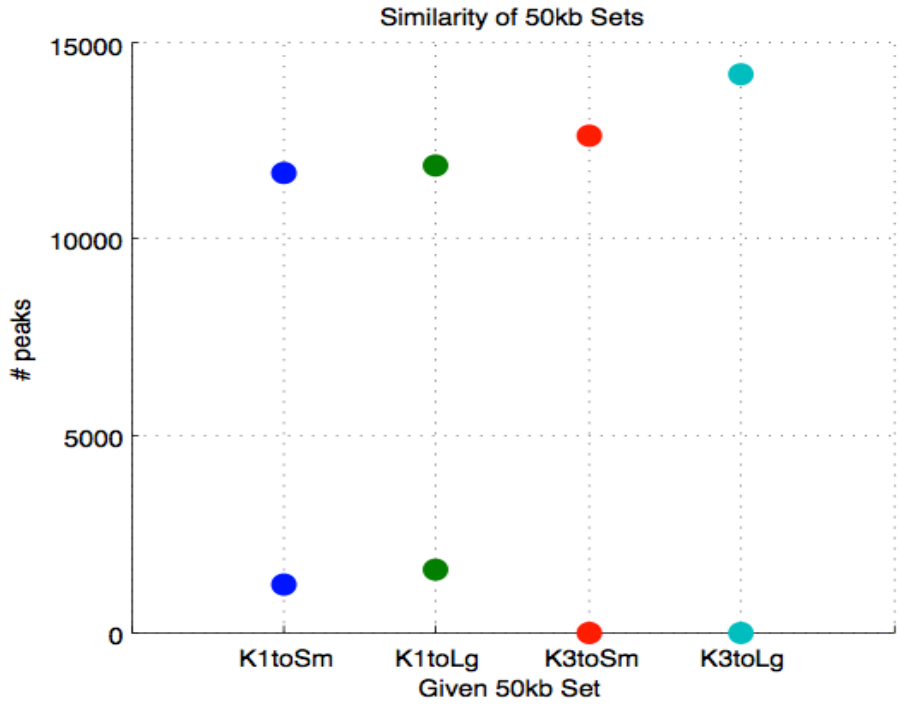


Figure 5. Similarity of peak sets within groups to the 50 kb set. The top set of data points represent the expected number of false peaks in each condition for the given

llocal value. The bottom set of data points represent how many peaks were in the smaller set of given llocal value that were NOT in the 50 kb set. The latter number is always a tiny fraction of total peaks (total peaks all >66,000; not shown here, see above figures). Moreover, it is always a small percentage of the number of peaks expected to be false suggesting that the discrepancies could be explained by differences in false peaks alone. If the number was greater than the number of expected false, then one would have to conclude that true peaks definitely differed between sets. Though some true peaks may differ here, even if all, that number is small. Therefore, the 50 kb sets were considered fine representatives of each condition.

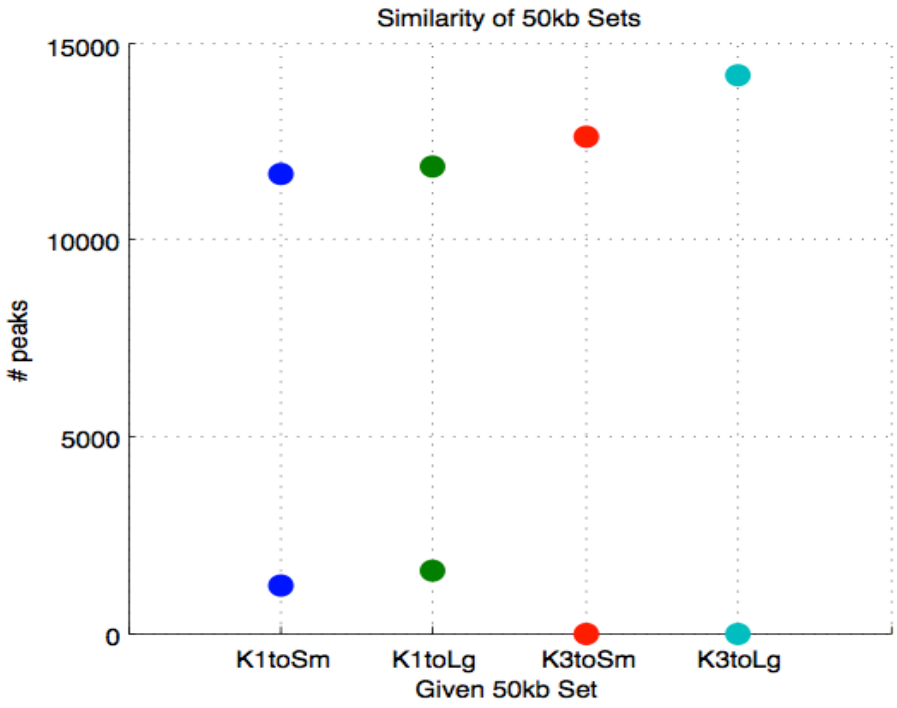


Figure 6. Similarity of the number of peaks regardless of the 50 kb set used. The top row of data points is expected number of false positives. The bottom row of data points are the number of peaks in the smaller set NOT in the largest K3toLg set. Both K1 sets had between 1000 and 2000 peaks that were not in the K3toLg set – numbers much less than the expected number of false peaks and far less than the total number of peaks. The K3toSm set is a proper subset of K3toLg. This is not surprising as all 'toSm' sets considered are proper subsets of their corresponding 'toLg' sets. All sets are therefore considered to be reasonably similar. As the K1 sets had lower FDR, one of these was chosen as our final set. Scaling to small is supposed to have higher

specificity and lower FDR. Nonetheless, we do not necessarily see this for the K1 sets. The K1toLg set actually seems to have a lower FDR.

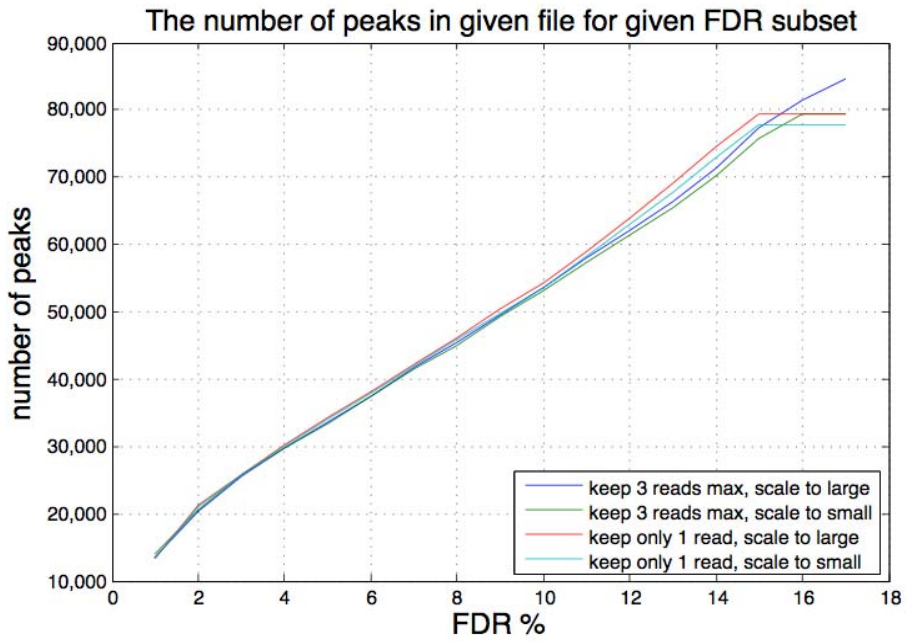


Figure 7. Effect of varying the False Discovery Rate on the number of peaks.

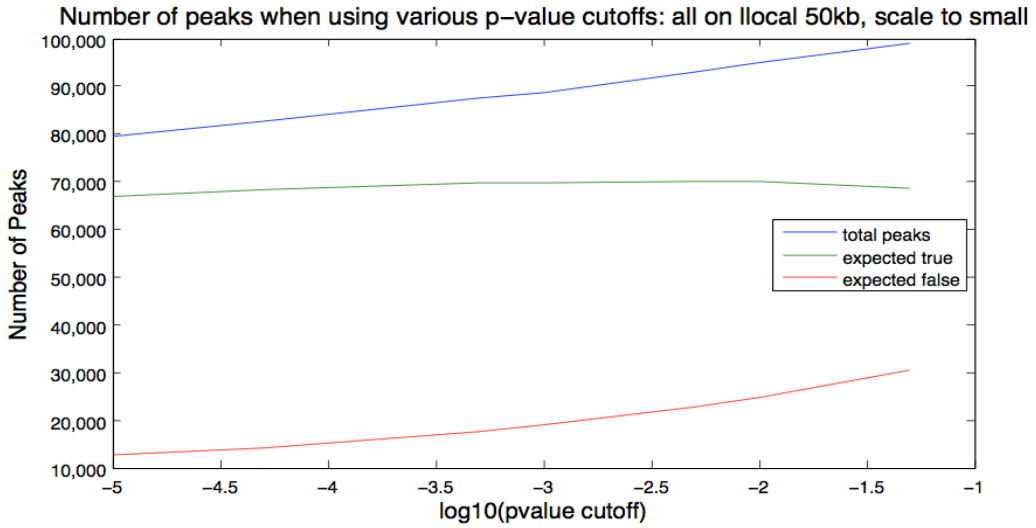


Figure 8. Effect of various p-value cutoffs on the number of peaks.

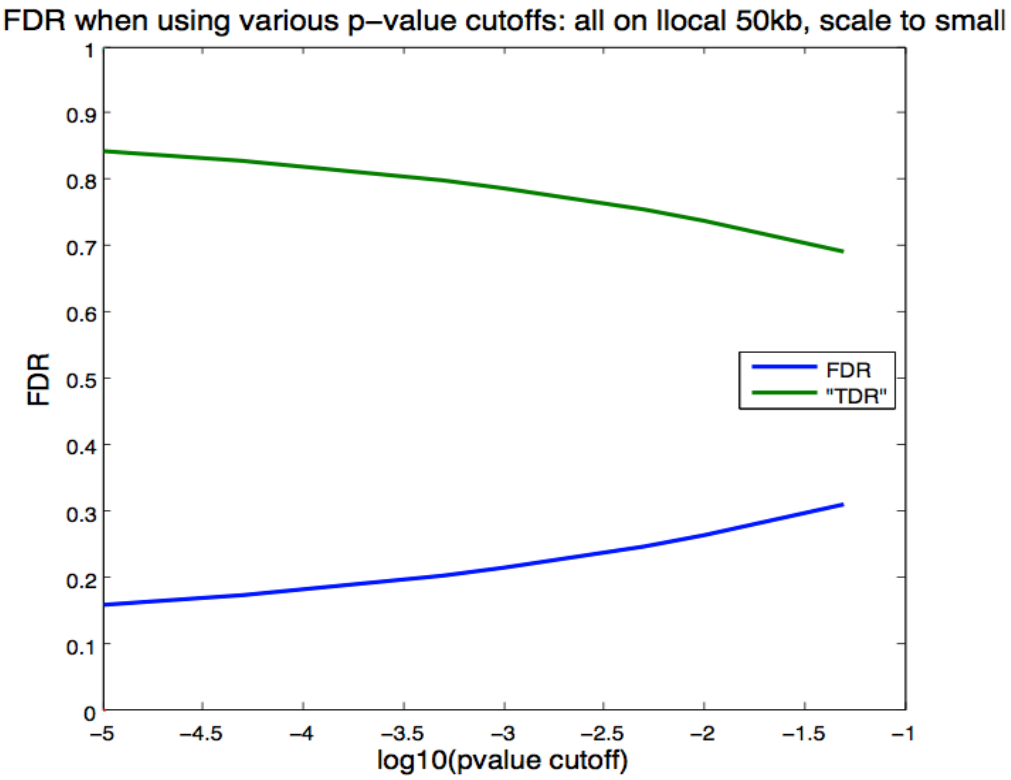


Figure 9. Effect of varying the p-value cutoff on the False Discovery Rate.

Biological variability between replicates...
All slocal=3kb, llocal=50kn, K1toLg
Rep1,2,3 corresponds to lanes 1,3,5 (sept 2011)

Note: Combined Aligned set also included reads from GAIIx run not included here

	NumReads	NumMapReads	NumMapK1Reads	NumMapK1ControlReads	NumPeaks
Rep1	128247879	54642570	43542420	43396363	80769
Rep2	139320824	84037740	68974326	68410750	55029
Rep3	136227515	89097527	65186902	64579951	62989
Combined Aligned	-	239583014	171804663	165182107	79173

	FDR
Rep1	4.4%
Rep2	12.28%
Rep3	13.1%
Combined Aligned	14.95%

Figure 10. Peaks called from three different samples of MCF-7 nascent DNA.

How many peaks in the ROW set are represented in the COLUMN set?				
	Rep1	**Rep2**	**Rep3**	**Combined Aligned**
Rep1	**80769**	33915	24808	53225
Rep2	32056	**55029**	40433	53591
Rep3	24115	39635	**62989**	53689
Combined Aligned	46013	49492	52875	**79339**

Figure 11. Shared peaks by the different samples of MCF-7 nascent DNA.

What % of the ROW set is represented in the COLUMN set?				
	Rep1	**Rep2**	**Rep3**	**Combined Aligned**
Rep1	100	41.99011997	30.71475442	65.89780733
Rep2	58.25292119	100	73.47580367	97.3868324
Rep3	38.28446237	62.92368509	100	85.23551731
Combined Aligned	57.9954373	62.38041821	66.64439935	100

Figure 12. Percent peak overlap in different samples of MCF-7 nascent DNA.

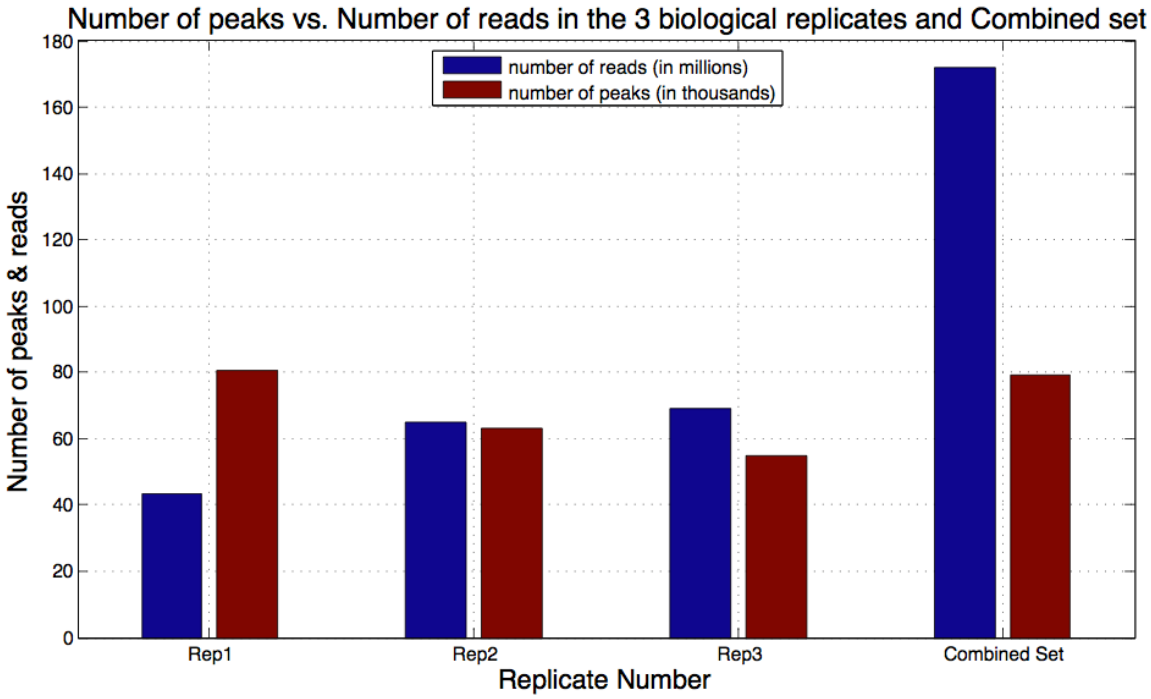


Figure 13. Number of peaks vs number of reads to suggest saturation.

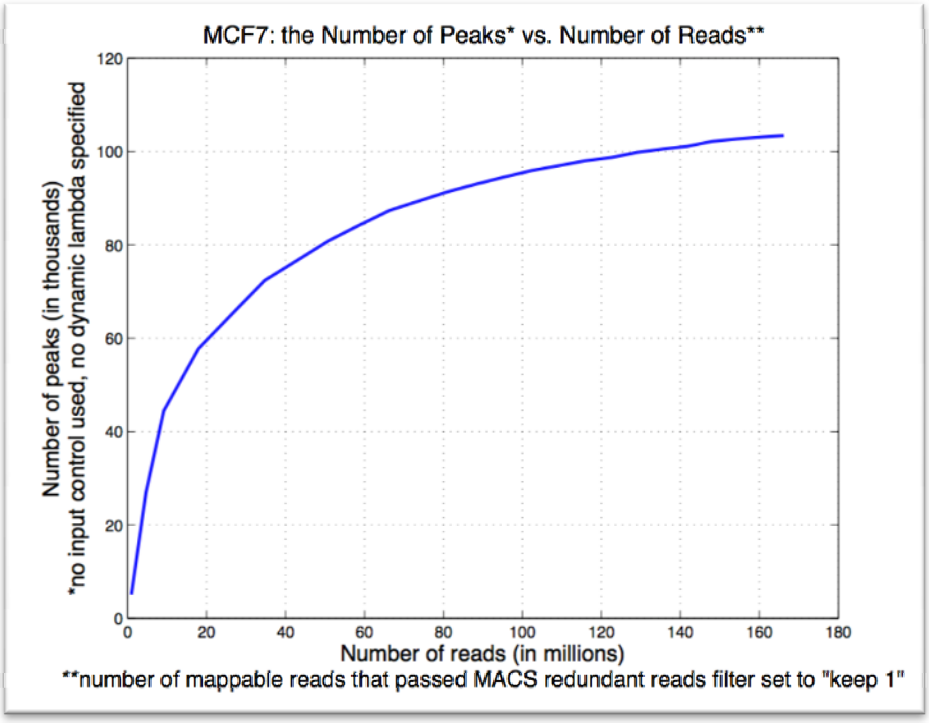


Figure 14. Saturation curve of number of peaks vs number of reads.

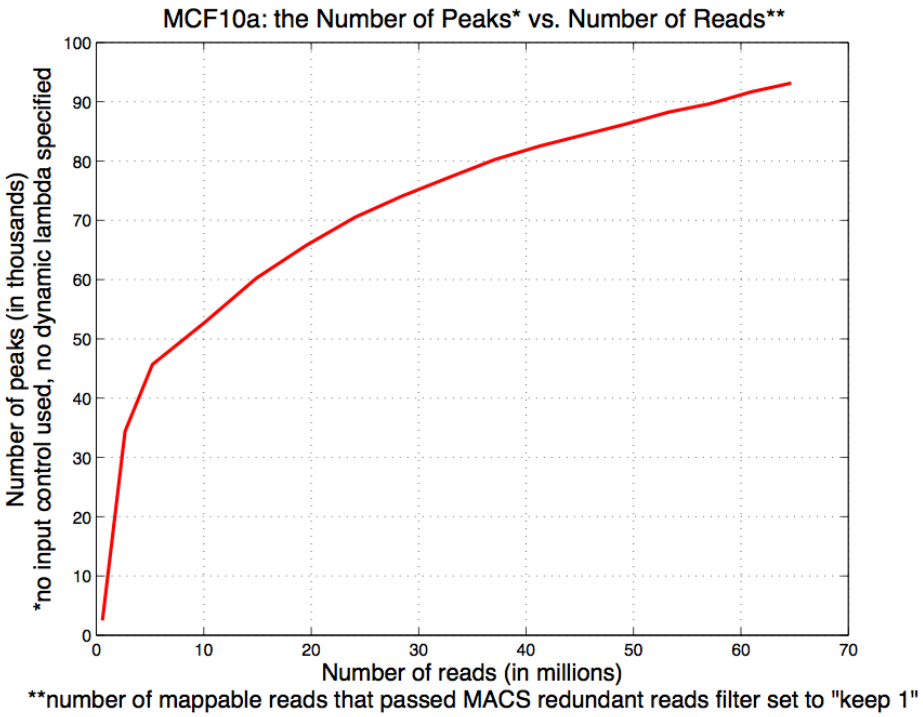


Figure 15. NS-Seq on MCF-10A replicating DNA approaches saturation.

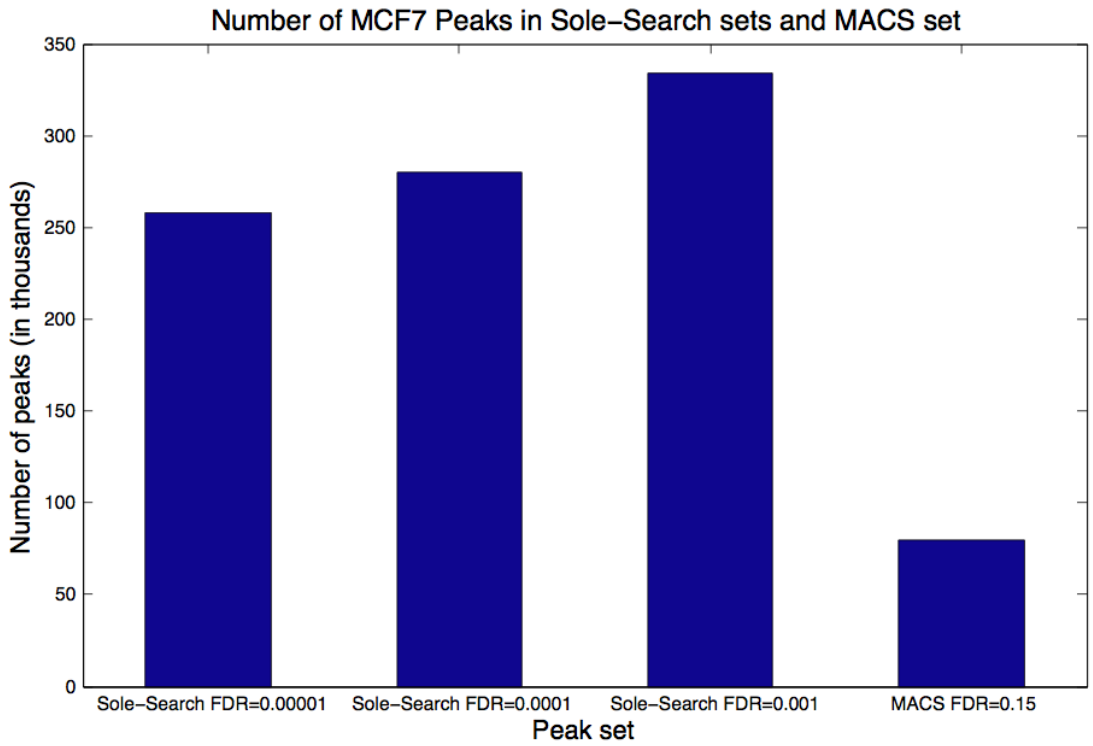


Figure 16. More MCF-7 peaks with Sole-Search than with MACS.

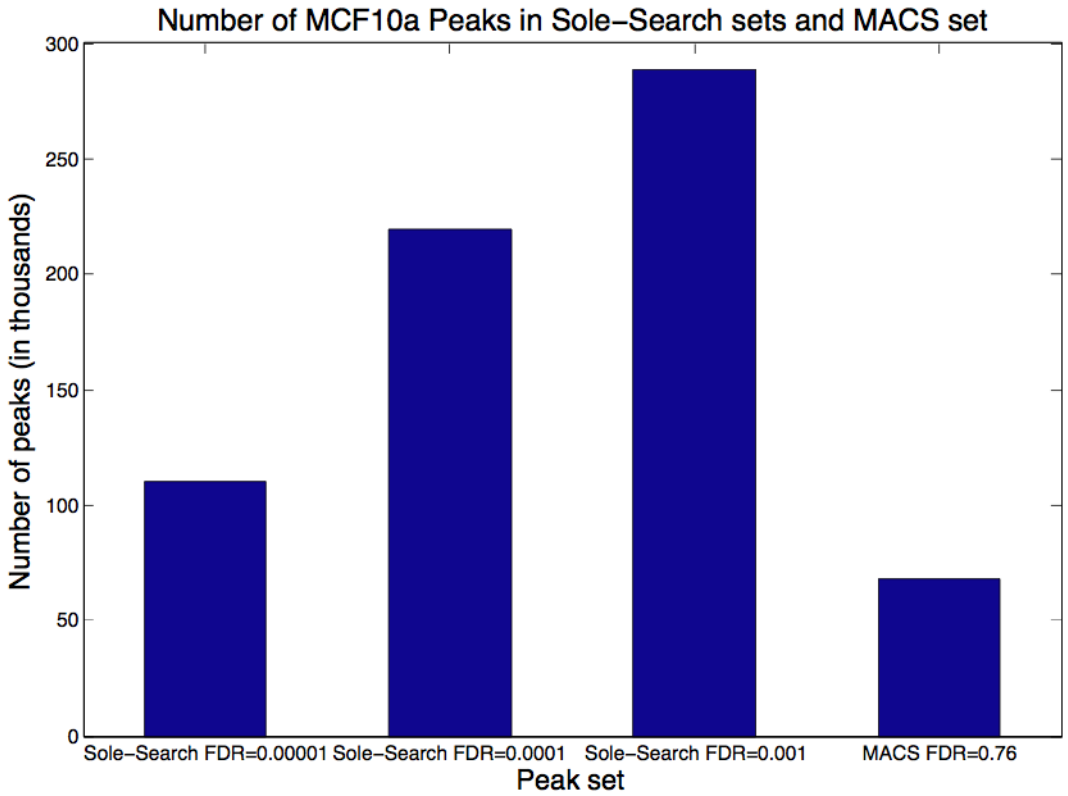


Figure 17. More MCF-10A peaks with Sole-Search than with MACS.

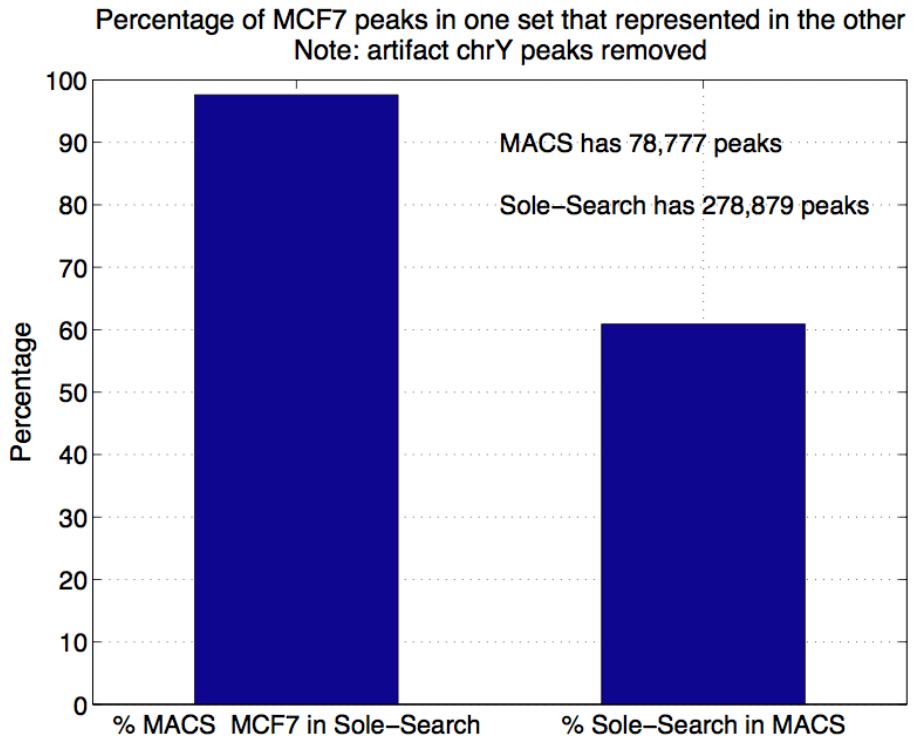


Figure 18. Percentage of MCF-7 peaks shared by MACS and Sole-Search.

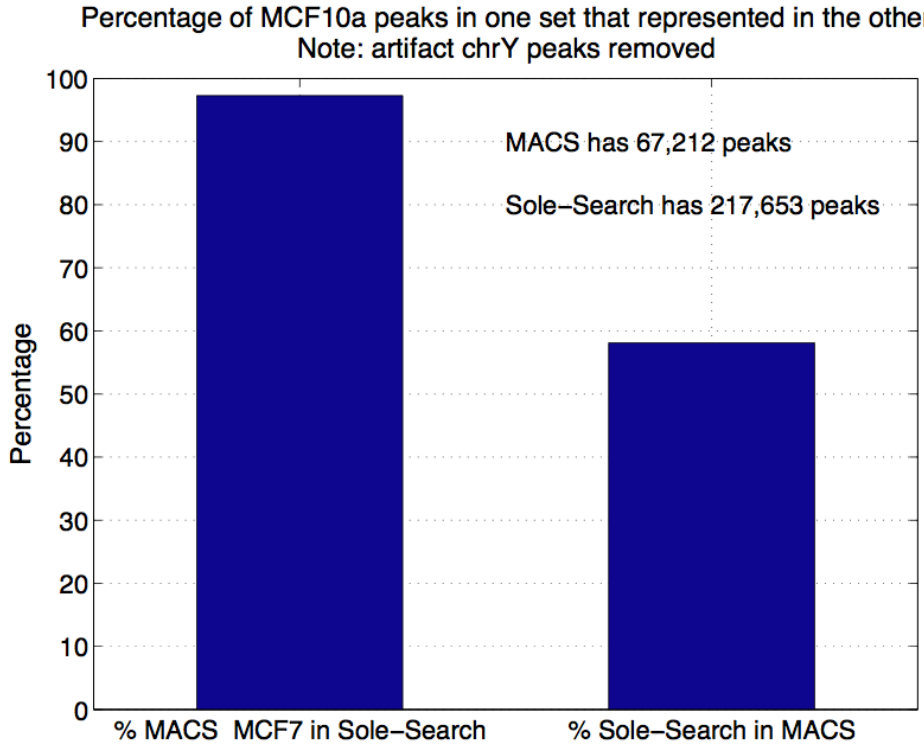


Figure 19. Percentage of MCF-10A peaks shared by MACS and Sole-Search.

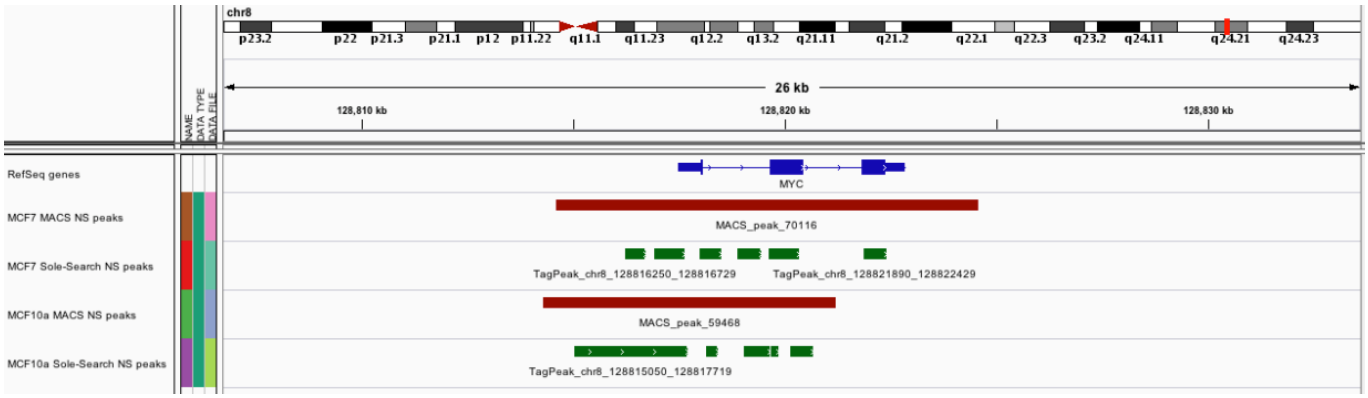


Figure 20. Peaks in the Myc locus called with MACS and with Sole-Search.

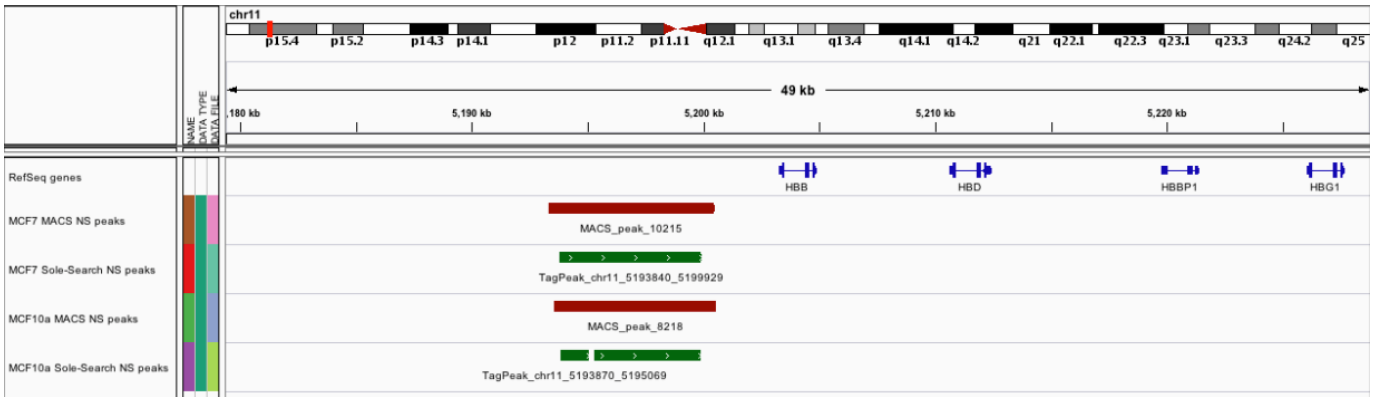


Figure 21. Peaks in the HBB locus called with MACS and with Sole-Search.

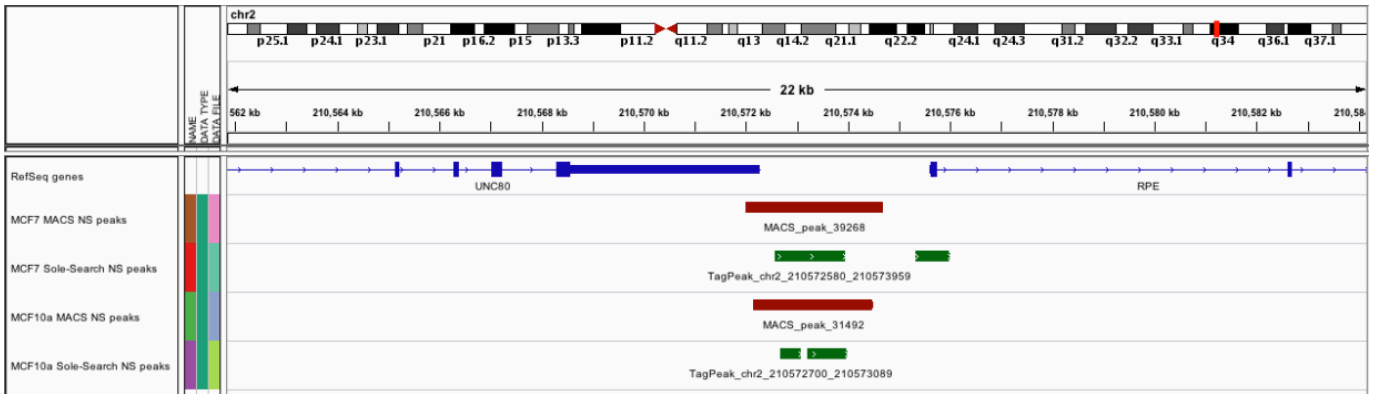


Figure 22. Peaks in the RPE locus called with MACS and with Sole-Search.

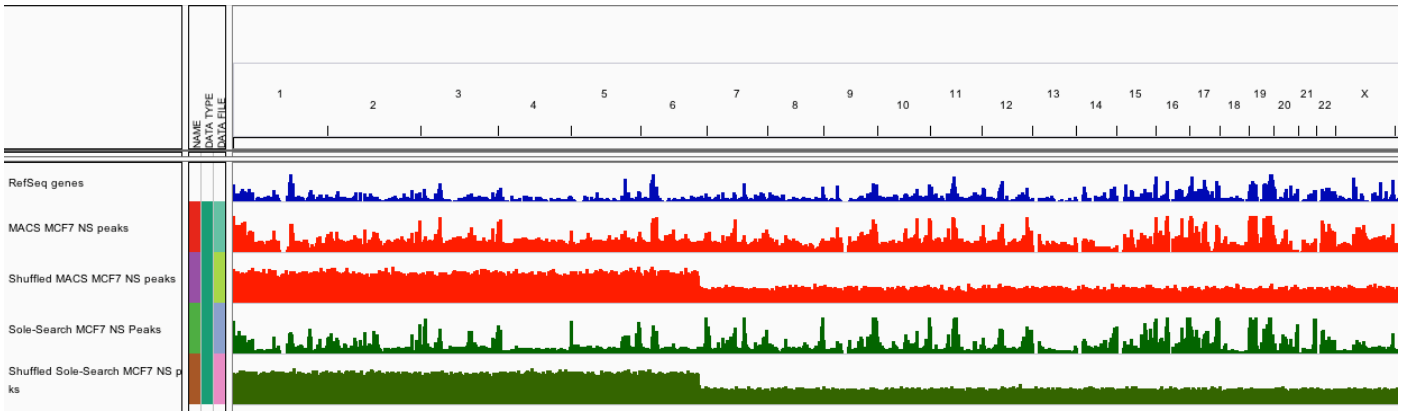


Figure 23. Density of peaks in the human genome called by MACS and by Sole-Search.

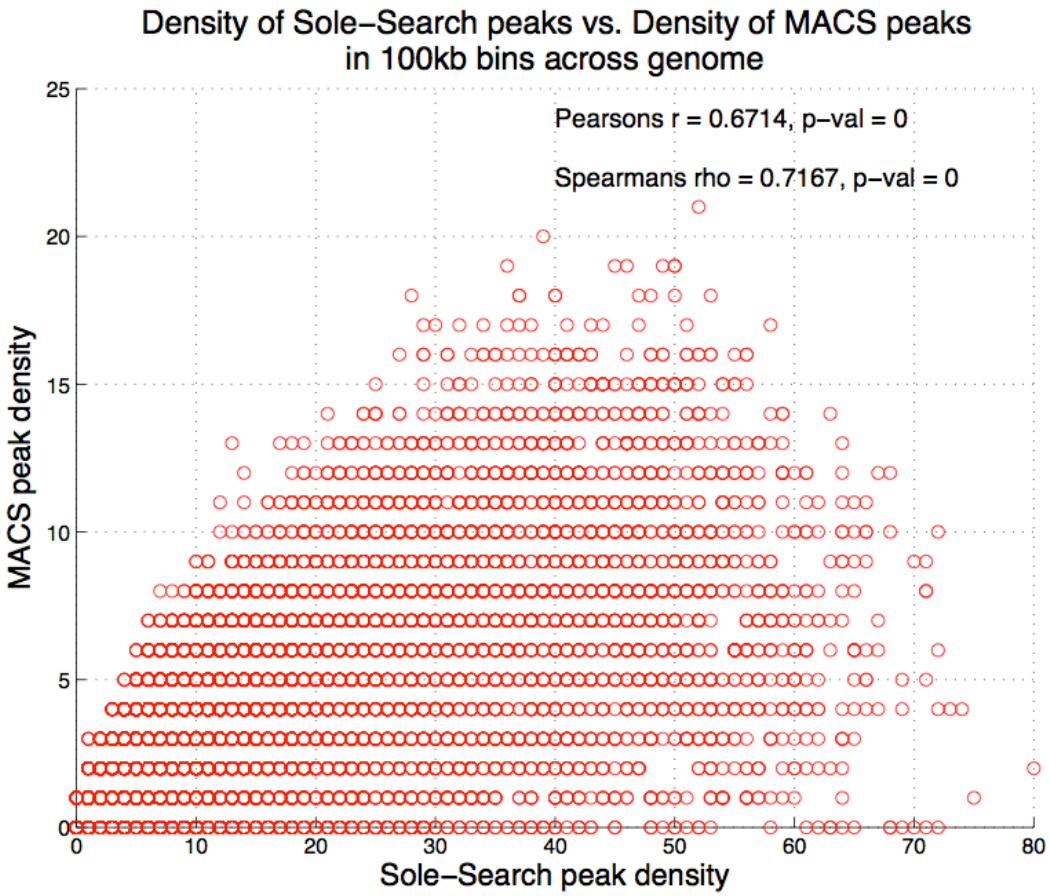


Figure 24. Density of Sole-Seach vs MACS peaks in 100 kb bins across the genome.

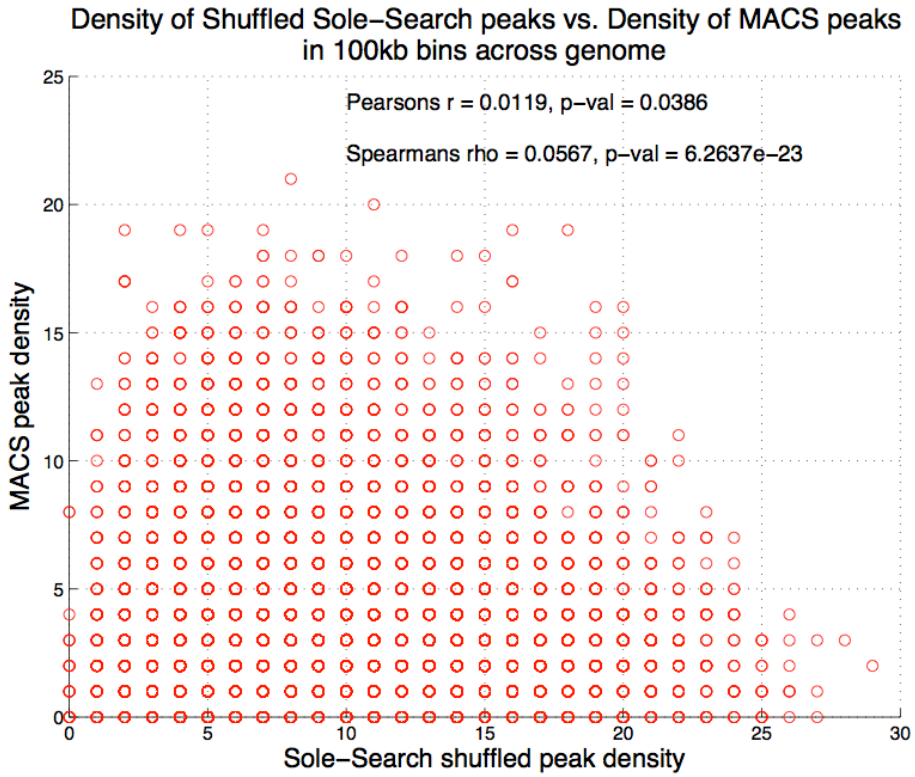


Figure 25. Density of Sole-Seach vs MACS shuffled peaks in 100 kb bins in the genome.

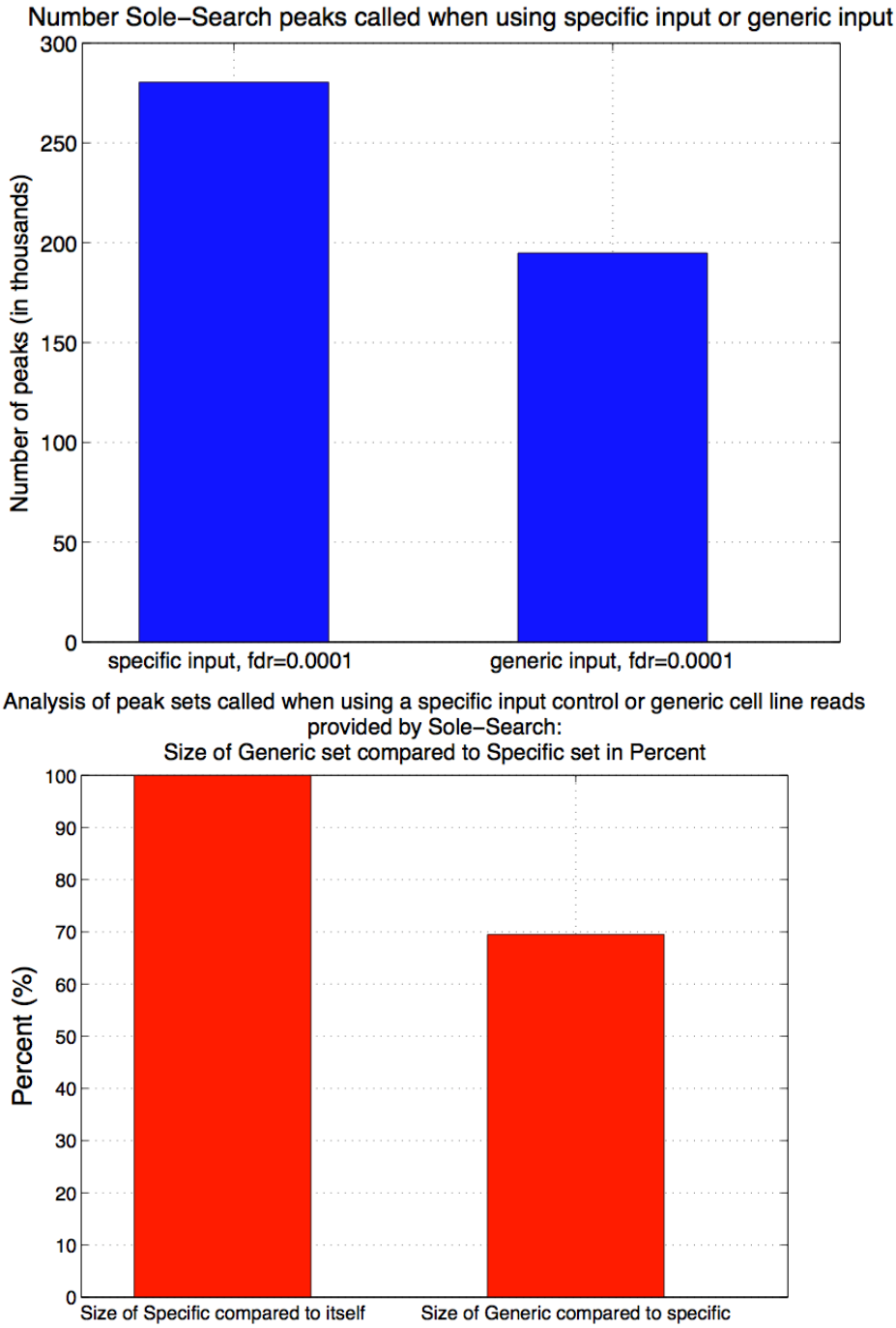


Figure 26. (a) Number or **(b)** percent of Sole-Search peaks using specific or generic input control.

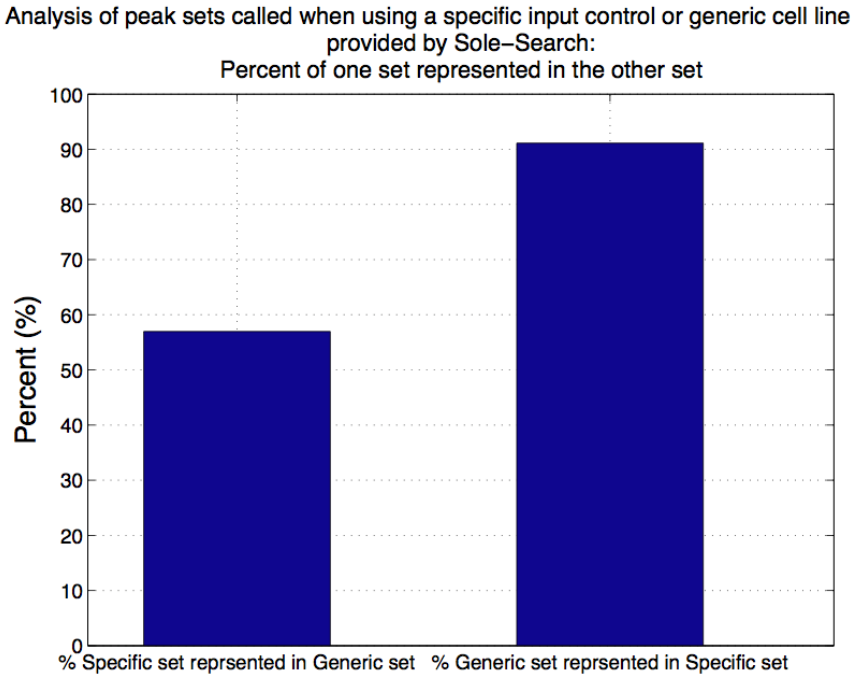


Figure 27. Percent specific set in generic set and vice versa when using a specific input or generic input control with Sole-Search.

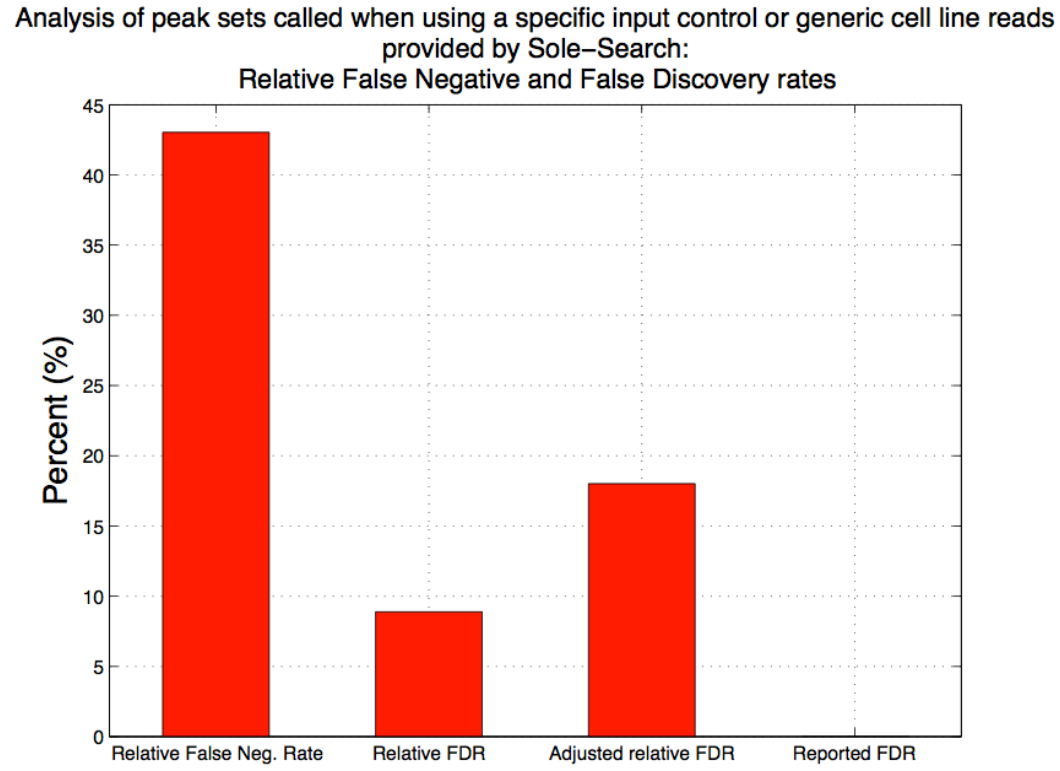


Figure 28. Relative False Negative and False Discovery Rates when using a specific input or generic input control with Sole-Search.

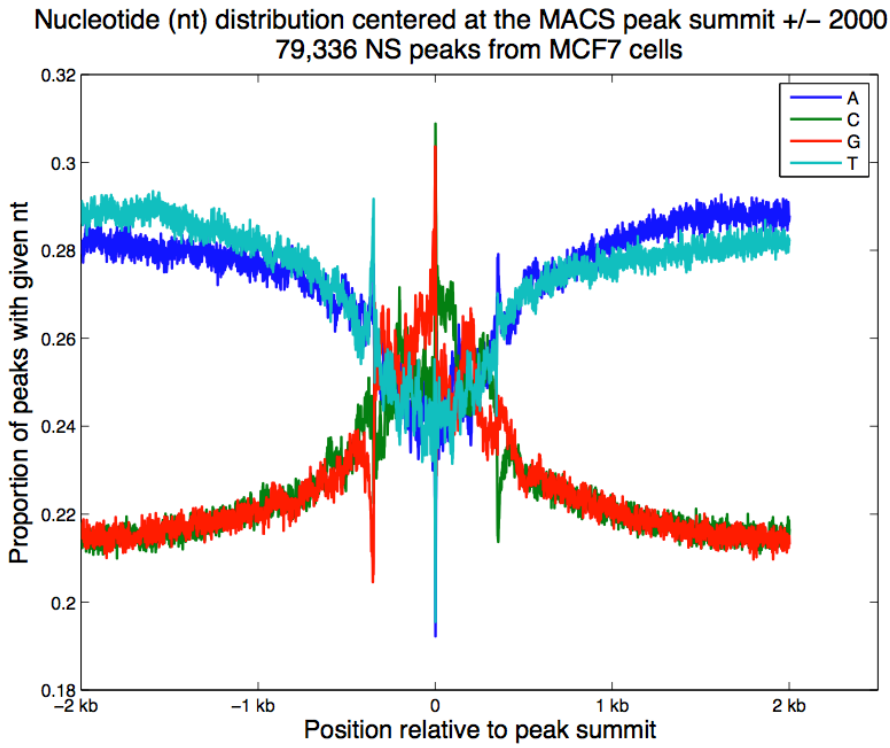


Figure 29. MACS MCF7 NS peak summits

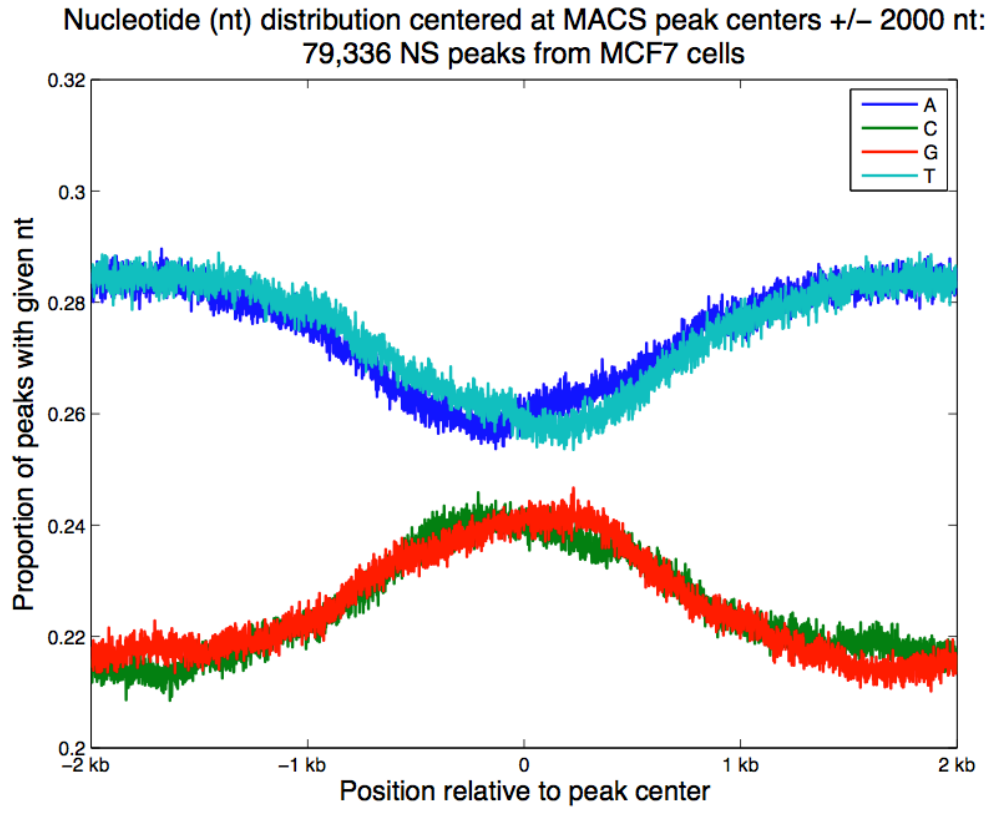


Figure 30. MACS MCF7 NS peak centers

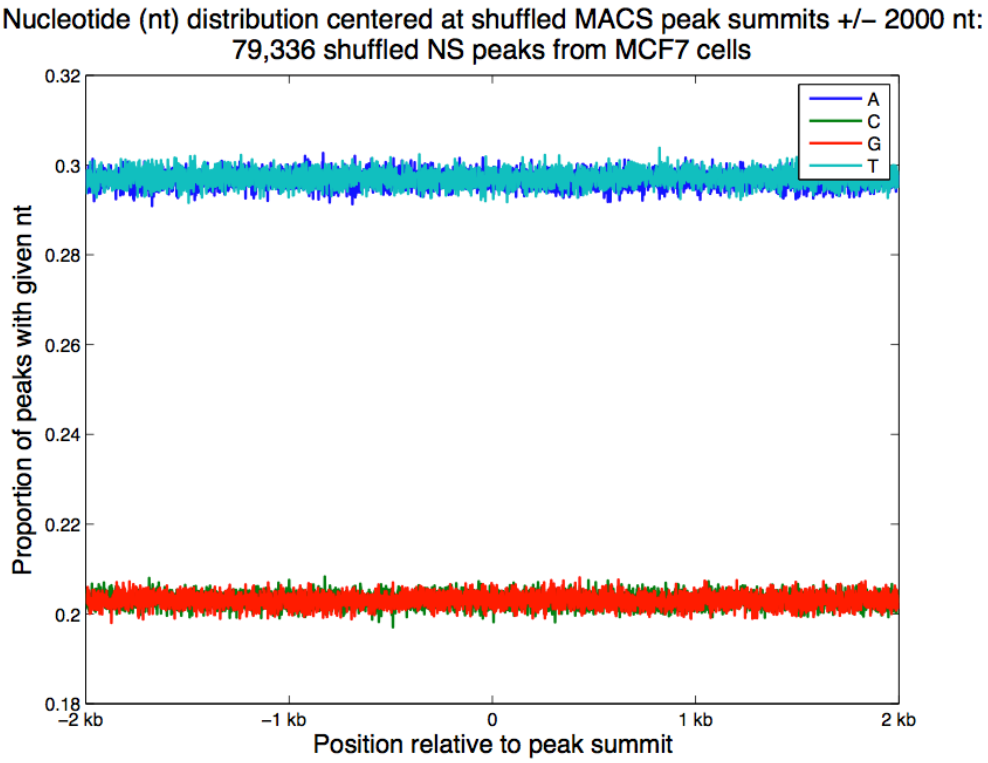


Figure 31. MACS MCF7 shuffled NS peak summits

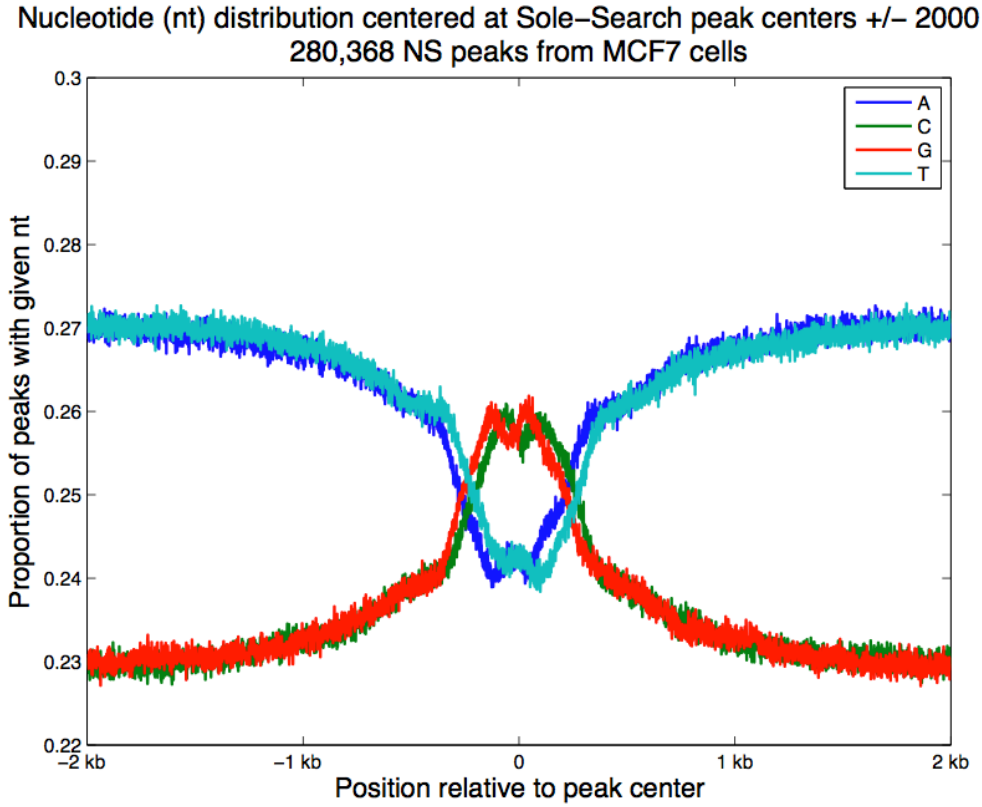


Figure 32. Sole-Search MCF7 NS peak centers

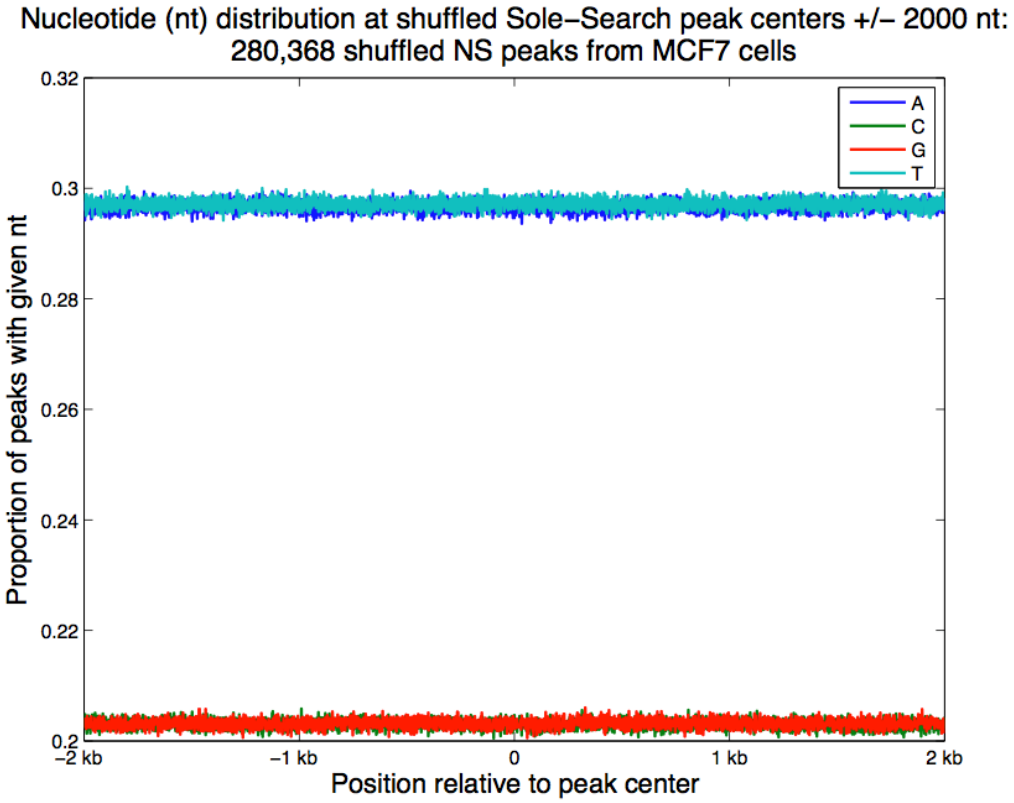


Figure 33. Sole-Search MCF7 shuffled NS peak centers

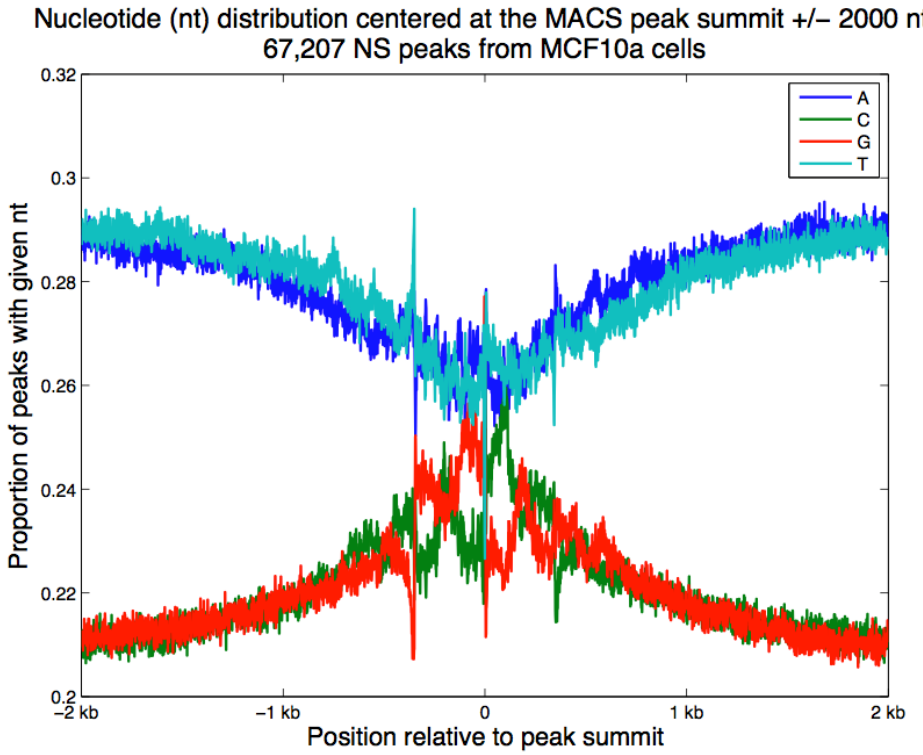


Figure 34. MACS MCF-10A NS peak summits

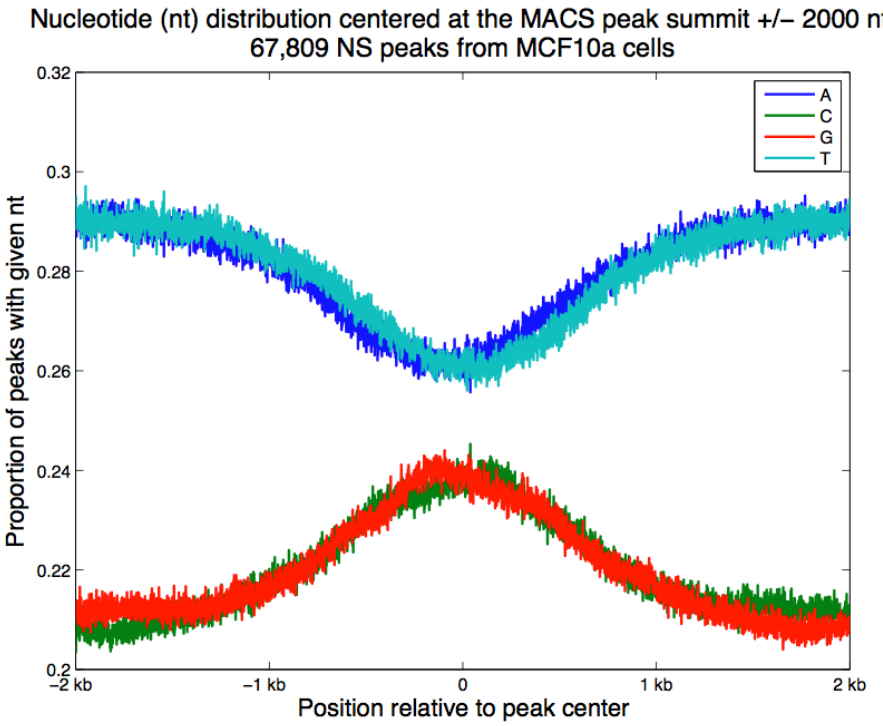


Figure 35. MACS MCF-10A peak centers

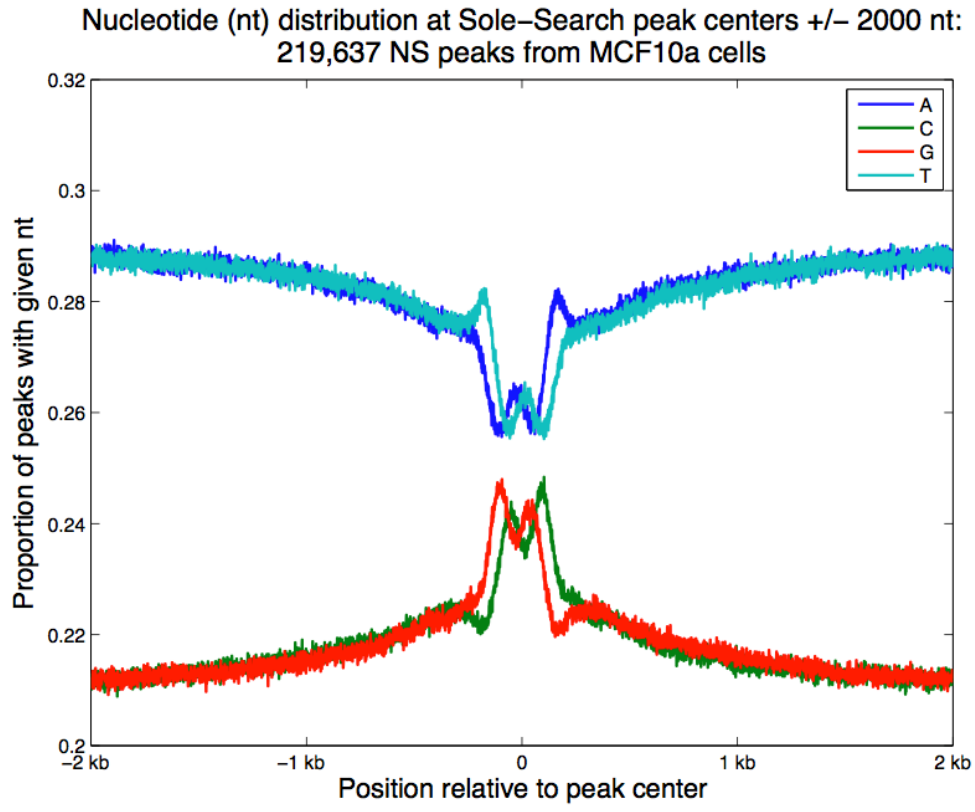


Figure 36. Sole-Search MCF-10A peak centers

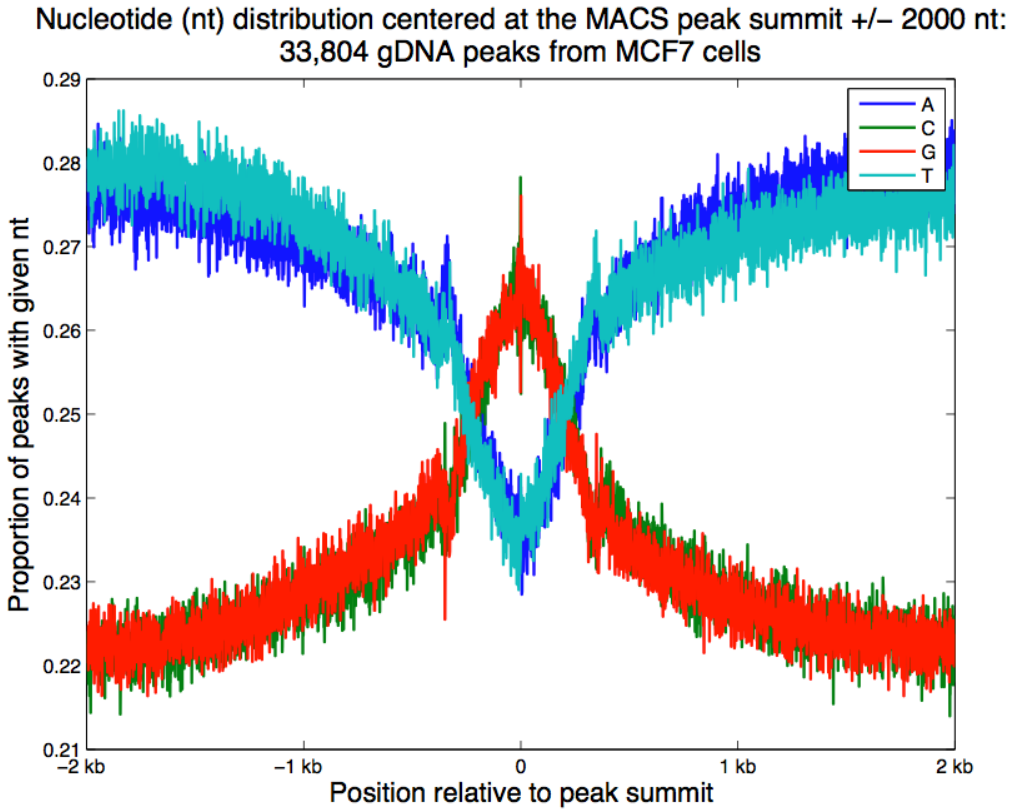


Figure 37. GC skew at peak summits from nonreplicating genomic DNA input

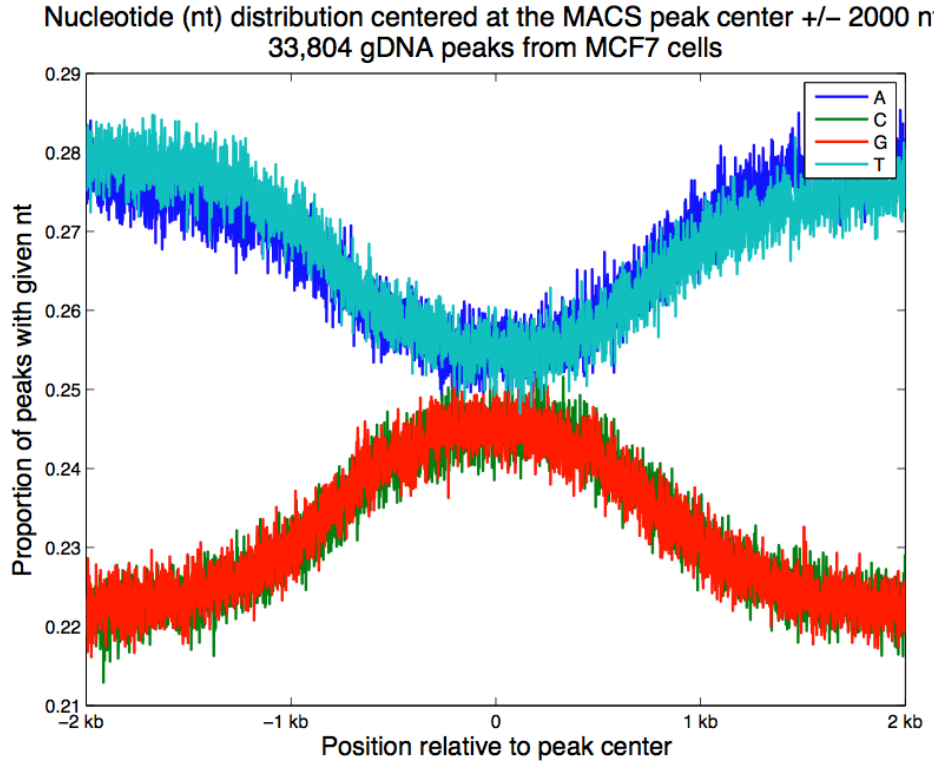


Figure 38. GC skew at peak centers from nonreplicating genomic DNA input

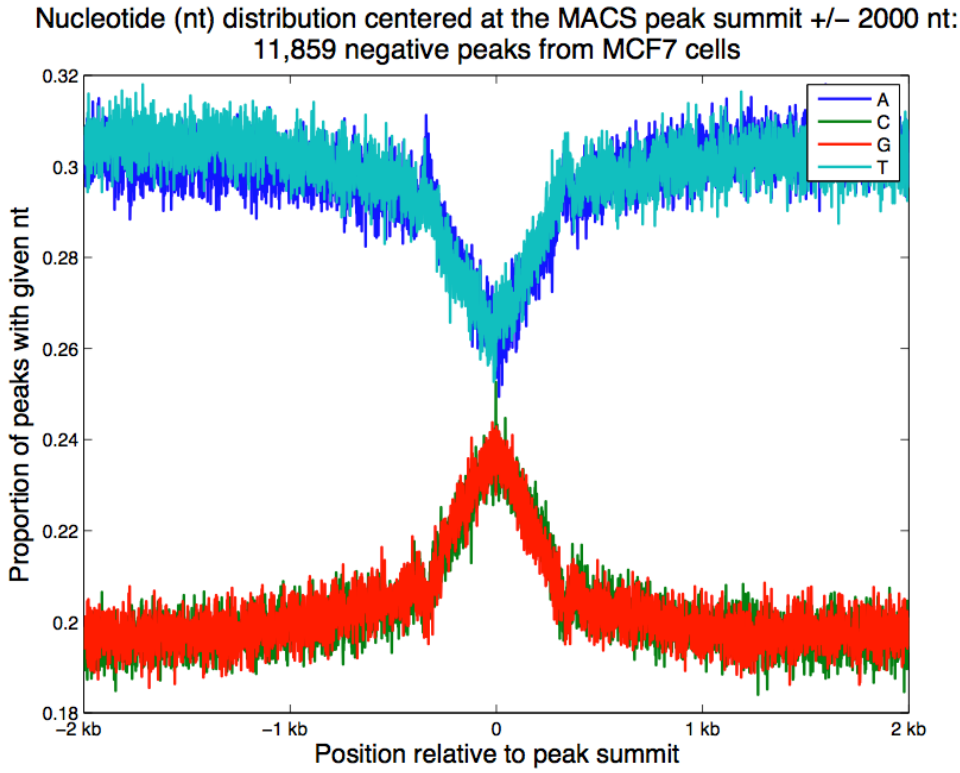


Figure 39. GC skew at peak summits from MCF-7 negative peaks

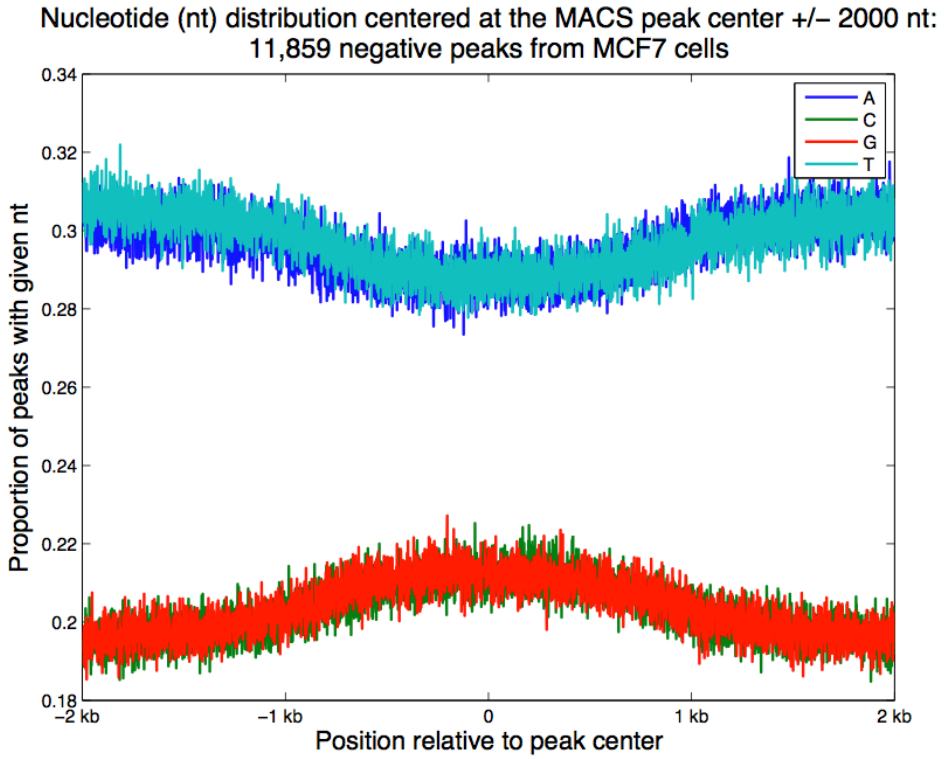


Figure 40. GC skew at peak centers from MCF-7 negative peaks

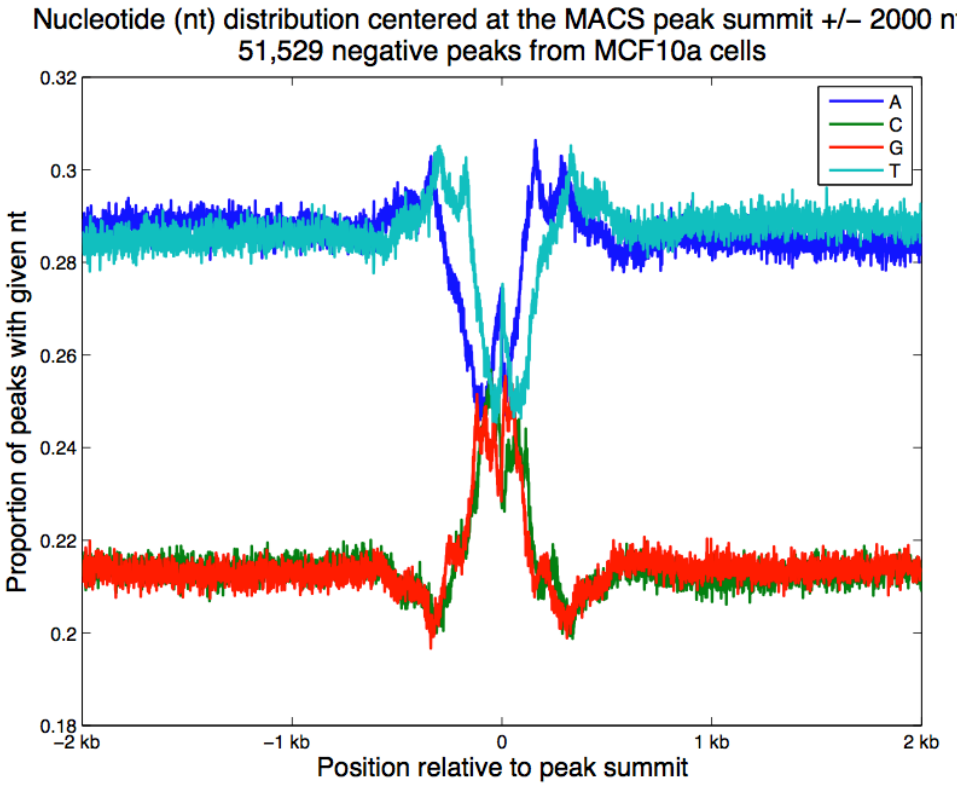


Figure 41. GC skew at peak summits from MCF-10A negative peaks

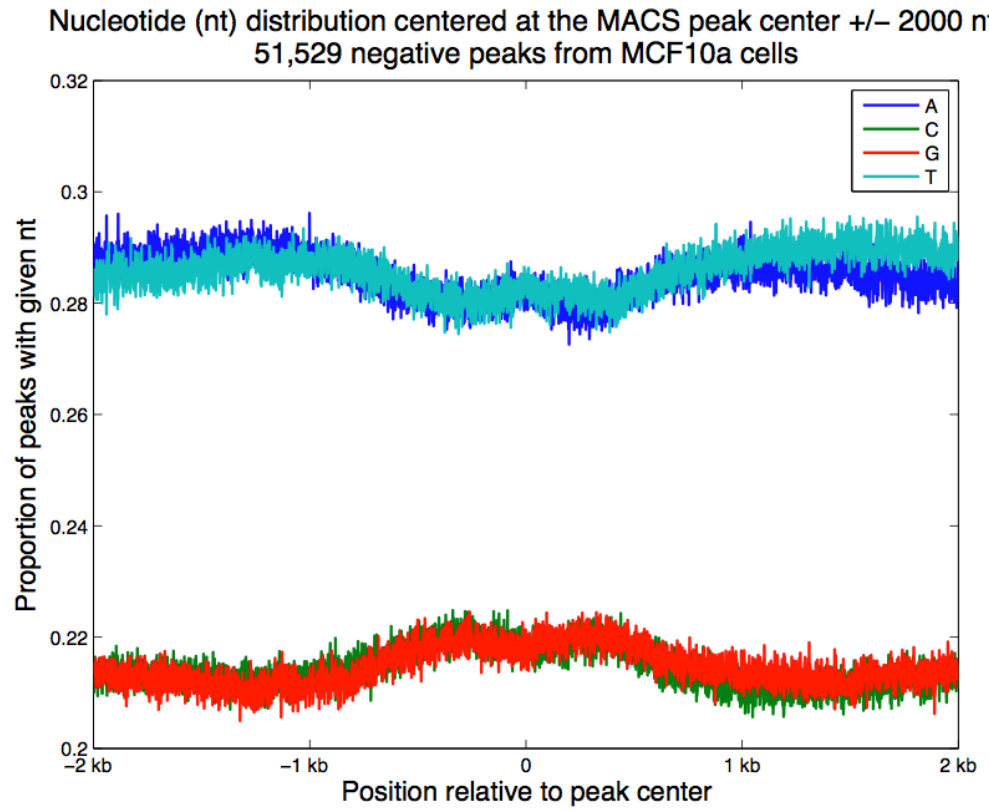


Figure 42. GC skew at peak centers from MCF-10A negative peaks

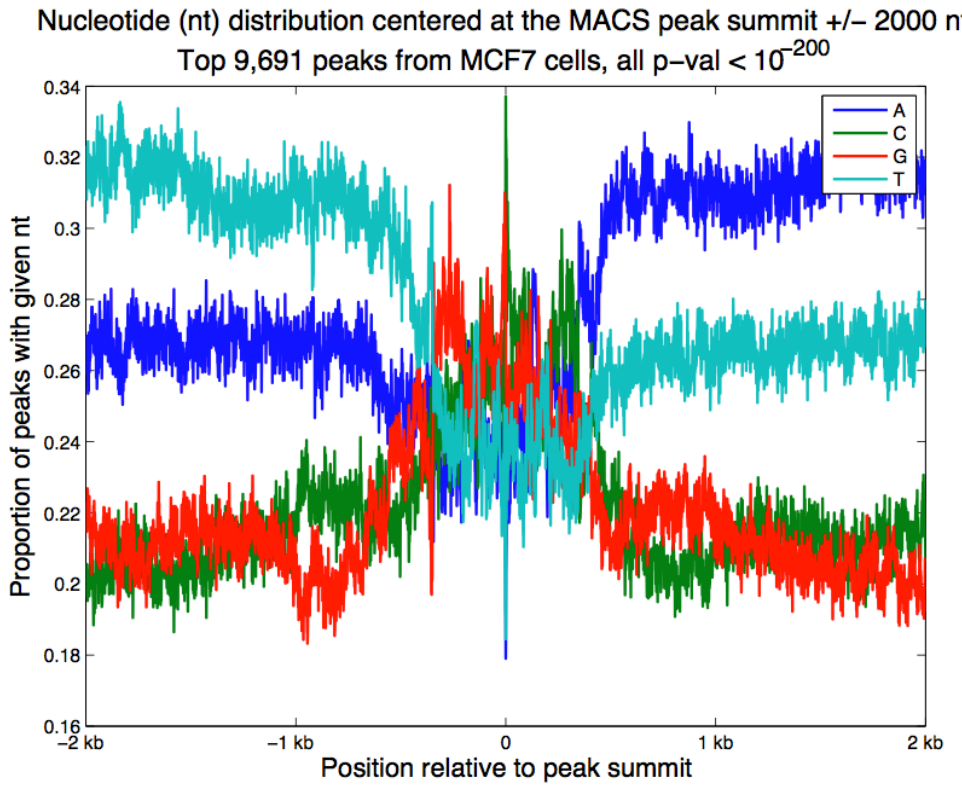


Figure 43. Base composition at peak summit of top 9691 MCF-7 Nascent Strand peaks.

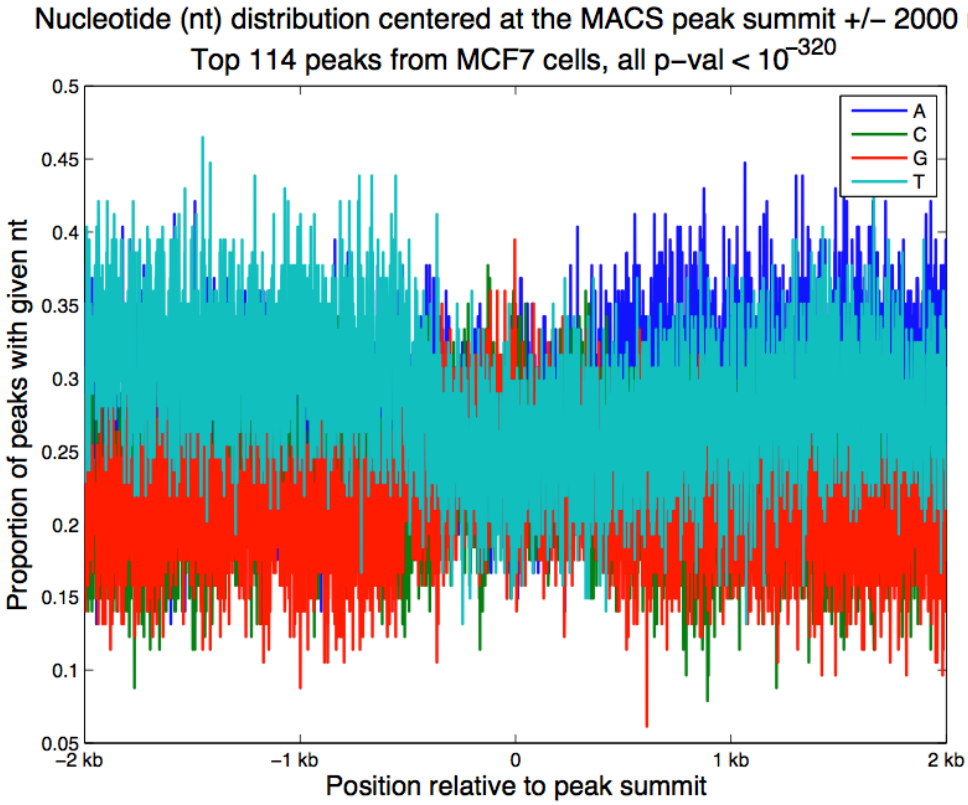


Figure 44. Base composition at peak summit of best 114 MCF-7 Nascent Strand peaks.